

Regarding Psychologists Testily: Legal Regulation of Psychological Assessment in the Public Schools

Donald N. Bersoff

Follow this and additional works at: <http://digitalcommons.law.umaryland.edu/mlr>



Part of the [Constitutional Law Commons](#)

Recommended Citation

Donald N. Bersoff, *Regarding Psychologists Testily: Legal Regulation of Psychological Assessment in the Public Schools*, 39 Md. L. Rev. 27 (1979)

Available at: <http://digitalcommons.law.umaryland.edu/mlr/vol39/iss1/5>

This Article is brought to you for free and open access by the Academic Journals at DigitalCommons@UM Carey Law. It has been accepted for inclusion in Maryland Law Review by an authorized administrator of DigitalCommons@UM Carey Law. For more information, please contact smccarty@law.umaryland.edu.

REGARDING PSYCHOLOGISTS TESTILY: LEGAL REGULATION OF PSYCHOLOGICAL ASSESSMENT IN THE PUBLIC SCHOOLS*

DONALD N. BERSOFF**

From its beginnings in World War I, the use of psychological tests has become so widespread that it is likely that every person in the United States has been affected in some way by the results of such tests.¹ Testing has become the means by which major decisions about people's lives are made in industry, hospitals, mental health clinics, the military, and, most pertinent here, the public schools. Tests themselves, by and large, are facially neutral. They do not inherently discriminate against those who take them and, undoubtedly, scores derived from tests have been used to admit, advance, and employ. For most people, however, test results have served as exclusionary mechanisms — to segregate, institutionalize, track, and deny access to desired goals. It is the unjustified negative consequences of psychological testing that have evoked legal concern, and it is the law's response that this Article explores.

Although testing affects virtually all Americans,² its most pervasive use is in the public schools. It has been estimated that more than 250 million standardized tests of academic ability, perceptual and motor skills, emotional and social characteristics, and vocational

* ©1980 Donald N. Bersoff. All rights reserved. A revised version of this Article will appear as a chapter in 3 *Perspectives in Law and Psychology: Testing and Evaluation* (B. Sales ed. 1980).

** Associate Professor, University of Maryland School of Law; Professor of Psychology, The Johns Hopkins University. B.A. 1958, M.A. 1960, Ph. D. 1965, New York University; J.D. 1976, Yale University.

1. Contrary to popular belief, however, it was not Binet, his enthusiastic American adopters (e.g., Terman, Goddard), or the World War I test developers (e.g., Otis) who initiated the testing movement, but the Chinese of 3000 years ago. "[T]hey invented the psychological test, applying it to government, the very framework of their society, in such manner that the test-makers, in effect, determined over many centuries much of the format of Chinese society." Dubois, *A Test-Dominated Society: China, 1115 B.C. — 1905 A.D.*, in TESTING PROBLEMS IN PERSPECTIVE 29 (A. Anastasi ed. 1966). For brief histories of the testing movement in this country, see Crissey, *Mental Retardation: Past, Present, and Future*, 30 AM. PSYCH. 800, 803-05 (1975); Cronbach, *Five Decades of Public Controversy Over Mental Testing*, 30 AM. PSYCH. 1 (1975); Laosa, *Nonbiased Assessment of Children's Abilities: Historical Antecedents and Current Issues*, in PSYCHOLOGICAL AND EDUCATIONAL ASSESSMENT OF MINORITY CHILDREN 1-6 (T. Oakland ed. 1977). For a black psychologist's perspective, see R. GUTHRIE, *EVEN THE RAT WAS WHITE: A HISTORICAL VIEW OF PSYCHOLOGY* (1976).

2. See D. GOSLIN, *THE SEARCH FOR ABILITY, passim* (1963).

interests and talent are used annually in education.³ Tests are used in conjunction with almost every major educational practice: screening, placement, program planning, program evaluation, and assessment of individual progress.⁴ But because tests have been used to exclude and segregate, they are alleged to have undermined "the American public school ideal promoted by educational reformers in the last century, whereby the school would serve as an object lesson in equality and brotherhood by drawing students from every social, economic and cultural background into the close association of the classroom."⁵ As a result, testing, the means by which it was hoped schools would most effectively and efficiently attain this ideal, has suffered heavy criticism. In the words of one commentator, the results of testing may "place an indelible stamp of intellectual status . . . on a child . . ., and possibly do irreparable harm to his self-esteem and his educational motivation."⁶

[Testing] may lead to a narrow conception of ability.

. . . .

It may emphasize individual competition and success rather than social cooperation, and thus conflict with the cultivation of democratic ideals of human equality. It may foster conformity rather than creativity. It may involve cultural bias. It may neglect important intangibles. It may, particularly in the case of personality testing, involve unwarranted and offensive invasions of privacy It may reward specious test-taking skills, or penalize the lack of it [*sic*].⁷

3. O. BRIM, D. GLASS, J. NEULINGER, I. FIRESTONE, & S. LERNER, *AMERICAN BELIEFS AND ATTITUDES ABOUT INTELLIGENCE* 1 (1969); M. HOLMAN & R. DOCTER, *EDUCATIONAL AND PSYCHOLOGICAL TESTING* 3-4 (1972).

4. Screening involves identification, usually through group tests, of those that may need special attention. Placement involves evaluation for entrance into special education programs or ability tracks, usually through the administration of individual assessment devices. When used for program planning, tests help determine specific curricula and intervention methods. Tests used for program evaluation assess the efficacy of instructional programs. When used to measure individual progress, they serve to monitor academic achievement in those educational programs. J. SALVIA & J. YSSELDYKE, *ASSESSMENT IN SPECIAL AND REMEDIAL EDUCATION* 14-16 (1978).

5. Sorgen, *Testing and Tracking in Public Schools*, 24 *HASTINGS L. J.* 1129, 1137 (1973) (citing H. MANN, *THE REPUBLIC AND THE SCHOOL; THE EDUCATION OF FREE MEN* 8, 32-33 (7th ed. 1957)).

6. Ebel, *The Social Consequences of Educational Testing*, in *TESTING PROBLEMS IN PERSPECTIVE* 18, 20 (A. Anastasi ed. 1966).

7. *Id.* Ebel also offers recommendations for averting these adverse consequences and lists briefly the consequences of failing to test. For supporting and contrary views, see A. ANASTASI, *PSYCHOLOGICAL TESTING* 45-64 (4th ed. 1976); H. BLACK, *THEY SHALL NOT PASS* (1963); L. CRONBACH, *ESSENTIALS OF PSYCHOLOGICAL TESTING*

Although criticism of testing from social, political, and psychological commentators spans six decades, only in the last fifteen years have legal scholars begun to examine its use.⁸ Two trends may explain this rather current interest. First, there has been an increased judicial scrutiny of educational practices, and, second, with desegregation of the schools ordered, psychological and educational tests have been criticized as devices used to hamper integration. Test results have been

298-307, 509-15 (3rd ed. 1970); M. GROSS, *THE BRAIN WATCHERS* (1962); B. HOFFMAN, *THE TYRANNY OF TESTING* (1962); L. KAMIN, *THE SCIENCE AND POLITICS OF IQ* (1974); G. ROBB, L. BERNADONI & R. JOHNSON, *ASSESSMENT OF INDIVIDUAL MENTAL ABILITY* 20-24 (1972); J. SATTLER, *ASSESSMENT OF CHILDREN'S INTELLIGENCE* (1974); W. WHYTE, *THE ORGANIZATION MAN* (1956); M. YOUNG, *THE RISE OF THE MERITOCRACY* (1958); Anastasi, *Psychology, Psychologists, and Psychological Testing*, 22 AM. PSYCH. 297 (1967); Bernal, *A Response to "Educational Uses of Tests With Disadvantaged Subjects,"* 30 AM. PSYCH. 93 (1975); Bersoff, *Silk Purses into Sow's Ears: The Decline of Psychological Testing and a Suggestion for its Redemption*, 28 AM. PSYCH. 892 (1973); Cleary, Humphreys, Kendrick & Wesman, *Educational Uses of Tests With Disadvantaged Students*, 30 AM. PSYCH. 15 (1975); Cronbach, *Five Decades of Public Controversy Over Mental Testing*, 30 AM. PSYCH. 1 (1975); Deutsch, Fishman, Kogan, North & Whiteman, *Guidelines for Testing Minority Group Children*, 20 J. SOC. ISSUES 129 (Supp. 1964); Jackson, *On the Report of the Ad hoc Committee on Educational Uses of Tests with Disadvantaged Students: Another Psychological View from the Association of Black Psychologists*, 30 AM. PSYCH. 88 (1975); Jorgensen, *IQ Tests and Their Educational Supporters*, 29 J. SOC. ISSUES 33 (1973); Kamin, *Social and Legal Consequences of I.Q. Tests as Classification Instruments: Some Warnings from Our Past*, 13 J. SCH. PSYCH. 317 (1975); Laosa, *Nonbiased Assessment of Children's Abilities*, in *PSYCHOLOGICAL AND EDUCATIONAL ASSESSMENT OF MINORITY CHILDREN* 10-11 (T. Oakland ed. 1977); Lippman, *The Abuse of the Tests*, 32 NEW REPUBLIC 297 (1922); McClelland, *Testing for Competence Rather Than for "Intelligence,"* 28 AM. PSYCH. 1 (1973); McNemar, *Lost: Our Intelligence? Why?*, 19 AM. PSYCH. 871 (1964); Mercer, *Psychological Assessment and the Rights of Children*, in 1 ISSUES IN THE CLASSIFICATION OF CHILDREN 130 (N. Hobbs ed. 1975); Reschly, *Nonbiased Assessment*, in *SCHOOL PSYCHOLOGY: PERSPECTIVES AND ISSUES*. 215, 230-35 (G. Phye & D. Reschly eds. 1979); Samuda, *Problems and Issues in Assessment of Minority Group Children*, in *MAINSTREAMING AND THE MINORITY CHILD* 65 (R. Jones ed. 1976); Williams, *Black Pride, Academic Relevance & Individual Achievement*, 2 COUNSELING PSYCH. 18 (1970). See generally *Testing and Public Policy*, 20 AM. PSYCH. 857 (special issue: M. Amrine ed. 1965); *Assessing Minority Group Children*, 11 J. SCH. PSYCH. 294 (special issue: B. Phillips ed. 1974). See also authorities cited note 118 *infra*.

8. CENTER FOR LAW AND EDUCATION, *CLASSIFICATION MATERIALS* (rev. ed. 1973) [hereinafter cited as *CLASSIFICATION MATERIALS*]; D. KIRP & M. YUDOF, *EDUCATIONAL POLICY AND THE LAW* (1974); Bersoff & Miller, *Ethical and Legal Issues in Behavioral Assessment*, in *DESCRIBING LEARNER CHARACTERISTICS OF HANDICAPPED CHILDREN AND YOUTH* 131 (D. Sabatino & T. Miller eds. 1979); Kirp, *Schools as Sorters: The Constitutional and Policy Implications of Student Classifications*, 121 U. PA. L. REV. 705 (1973); Ruch & Ash, *Comments on Psychological Testing*, 69 COLUM. L. REV. 608, 610, 615 (1969); Shea, *An Educational Perspective of the Legality of Intelligence Testing and Ability Grouping*, 6 J.L. & EDUC. 137 (1977); Sorgen, *Testing and Tracking in Public Schools*, 24 HASTINGS L.J. 1129 (1973); Yudof, *Equal Educational Opportunity and the Courts*, 51 TEX. L. REV. 411 (1973); Note,

seen as tools of discrimination denying full realization of the constitutional rights of racial and ethnic minorities.⁹ As a result, since the mid-1960's there has been much litigation and legislation affecting the administration, interpretation, and use of psychological tests.

This Article will examine psychological testing in the public schools. It will survey recent judicial examinations of educational practices, consider the role of psychological testing in efforts to block desegregation, and evaluate current legislation concerning the testing process. The Article will then examine the implications for personality testing of the test subjects' privacy rights, and finally, discuss how

Legal Implications of the Use of Standardized Ability Tests in Employment and Education, 68 COLUM. L. REV. 691 (1968) [hereinafter cited as *Legal Implications*]; Comment, *Ability Grouping in Public Schools: A Threat to Equal Protection?*, 1 CONN. L. REV. 150 (1968) [hereinafter cited as *Ability Grouping*]; Note, *The Legal Implications of Cultural Bias in the Intelligence Testing of Disadvantaged School Children*, 61 GEO. L.J. 1027 (1973) [hereinafter cited as *Cultural Bias*]; Note, *Hobson v. Hansen: Judicial Supervision of the Color-Blind School Board*, 81 HARV. L. REV. 1511 (1968) [hereinafter cited as *Judicial Supervision*]; Note, *Segregation of Poor and Minority Children into Classes for the Mentally Retarded by the Use of IQ Tests*, 71 MICH. L. REV. 1212 (1973) [hereinafter cited as *Segregation*]; Comment, *Equal Protection and Standardized Testing*, 44 MISS. L.J. 900 (1973) [hereinafter cited as *Equal Protection*]; Comment, *Constitutional Law — Limiting the Use of Standardized Intelligence Tests for Ability Grouping in Public Schools*, 51 N. C. L. REV. 1564 (1973); Note, *Equal Protection and Intelligence Classifications*, 26 STAN. L. REV. 647 (1974) [hereinafter cited as *Intelligence Classifications*]; Note, *Educational Testing: A Challenge for the Courts*, 1973 U. ILL. L.F. 375 [hereinafter cited as *Educational Testing*]; Note, *Constitutional Requirements for Standardized Ability Tests Used in Education*, 26 VAND. L. REV. 789 (1973). See generally Horowitz, *Unseparate But Unequal — The Emerging Fourteenth Amendment Issue in Public School Education*, 13 U.C.L.A. L. REV. 1147 (1966); Kirp, *Law, Politics, and Equal Educational Opportunity: The Limits of Judicial Involvement*, 47 HARV. EDUC. REV. 117 (1977); McClung, *Competency Testing: Potential for Discrimination*, 11 CLEARINGHOUSE REV. 439 (1977); Teitelbaum & Hiller, *Bilingual Education: The Legal Mandate*, 47 HARV. EDUC. REV. 138 (1977). For a discussion of psychologists' perspectives on the law's impact, see Oakland & Laosa, *Professional, Legislative, and Judicial Influences on Psychoeducational Assessment Practices in Schools*, in *PSYCHOLOGICAL AND EDUCATIONAL ASSESSMENT OF MINORITY CHILDREN* 21 (T. Oakland ed. 1977).

9. "Such devices are now being used by deceptive, inept, or simply indifferent school boards to smokescreen the daily violation of individual rights." *Cultural Bias*, *supra* note 8, at 1040. There appears to be no case law and little discussion of the role tests play in sex discrimination among students in the public schools. Sex discrimination is barred by 20 U.S.C. § 1681 (1977). Implementing regulations to that statute forbid a recipient of federal funds to:

[A]dminister or operate any test or other criterion for admission which has a disproportionately adverse effect on persons on the basis of sex unless the use of such test or criterion is shown to predict success accurately and alternative tests or criteria which do not have such a disproportionately adverse effect are shown to be unavailable.

45 C.F.R. § 86.21 (1978). See also Flaughner, *The Many Definitions of Test Bias*, 33 AM. PSYCH. 671, 674 (1978).

psychologists may be made more accountable through access to records.

JUDICIAL SCRUTINY OF THE PUBLIC SCHOOLS

Because the use of tests to identify, evaluate, and place children is a time-honored educational prerogative, authorized by school boards and performed by school personnel, "[a] threshold issue is whether courts should involve themselves in [such] matters at all."¹⁰ There was a time when the behavior of school officials went virtually unexamined by the courts. Pleading lack of expert knowledge, judges were wary of interfering with the discretion of administrators to educate their students. For example, sustaining the exclusion of an allegedly mentally retarded child in 1893, a Massachusetts court stated:

The management of the schools involves many details; and it is important that a board of public officers . . . having jurisdiction to regulate the internal affairs of the schools, should not be interfered with or have their conduct called into question before another tribunal. . . . A jury composed of men of no special fitness to decide educational questions should not be permitted to say that their answer is wrong.¹¹

But while the 1954 landmark desegregation decision, *Brown v. Board of Education*,¹² "significantly altered this allocation of authority, [it did so] only where issues of race were concerned"¹³ A decade ago the Supreme Court was still warning:

[J]udicial interposition in the operation of the public school system of the Nation raises problems requiring care and restraint Courts do not and cannot intervene in the resolution of conflicts which arise in the daily operation of school systems and which do not directly and sharply implicate basic constitutional values.¹⁴

However, the past ten years have seen a marked increase in judicial involvement in schools. Educators' total immunity from

10. *Segregation*, *supra* note 8, at 1225.

11. *Watson v. City of Cambridge*, 157 Mass. 561, 563, 32 N.E. 864, 864-65 (1893). See also *State ex rel. Beattie v. Board of Educ.*, 169 Wis. 231, 233, 172 N.W. 153, 154 (1919) (upholding exclusion of physically handicapped child who allegedly had "a depressing and nauseating effect upon the teachers and school children.").

12. 347 U.S. 483 (1954).

13. *Kirp, Law, Politics, and Equal Educational Opportunity*, *supra* note 8, at 118.

14. *Epperson v. Arkansas*, 393 U.S. 97, 104 (1968).

judicial scrutiny may be said to have ended with the Supreme Court's 1969 declaration that "students in school as well as out of schools are 'persons' under our Constitution . . . possessed of fundamental rights which the State must respect."¹⁵ Since then the Supreme Court has decided such issues as the reach of compulsory education laws,¹⁶ the requirements of due process prior to the infliction of disciplinary¹⁷ and academic sanctions,¹⁸ the immunity of school officials from money damage liability for violations of students' civil rights,¹⁹ the allocation of financial resources to pupils in poor school districts,²⁰ the education of non-English speaking children,²¹ the permissibility of sex-separate high schools,²² the legality of special admissions programs for minorities,²³ the obligation of colleges and universities to admit handicapped students,²⁴ and, most recently, the validity of system-wide remedies to reduce school segregation.²⁵

15. *Tinker v. Des Moines Ind. Community School Dist.*, 393 U.S. 503, 511 (1969) (upholding the first amendment right of students to wear armbands in demonstrations against the Vietnam War).

16. *Wisconsin v. Yoder*, 406 U.S. 205 (1972) (upholding religious sect's claim that compulsory education requirement endangered first amendment right to free exercise of religion).

17. *Compare Ingraham v. Wright*, 430 U.S. 651 (1977) (public school authorities have right to impose corporal punishment on disruptive school children without prior hearing) with *Goss v. Lopez*, 419 U.S. 565 (1975) (student has right to a hearing prior to short term suspension).

18. *Board of Curators of Univ. of Mo. v. Horowitz*, 435 U.S. 78 (1978) (university authorities have right to dismiss student for "academic" causes without prior hearing when student has been informed of faculty dissatisfaction with her performance).

19. *Wood v. Strickland*, 420 U.S. 308, 322 (1975) ("A school board member is not immune from liability for damages under § 1983 if he knew or reasonably should have known that the action he took . . . would violate the constitutional rights of the student affected, or if he took the action with the malicious intention to cause a deprivation of constitutional rights or other injury to the student.").

20. *San Antonio Ind. School Dist. v. Rodriguez*, 411 U.S. 1 (1973) (upholding constitutionality of state's method of financing public education in proportion to amount of taxes paid, where system led to disparities in funding among school districts).

21. *Lau v. Nichols*, 414 U.S. 563 (1974) (non-English-speaking school children entitled to English language instruction under 1964 Civil Rights Act).

22. *Vorchheimer v. School Dist.*, 430 U.S. 703 (1977), *aff'g* by an equally divided court 532 F.2d 880 (3d Cir. 1976).

23. *Regents of Univ. of Cal. v. Bakke*, 438 U.S. 265 (1978).

24. *Southeastern Community College v. Davis*, 442 U.S. 397 (1979).

25. *Columbus Bd. of Educ. v. Penick*, 99 S. Ct. 2941 (1979); *Dayton Bd. of Educ. v. Brinkman*, 99 S. Ct. 2971 (1979) (*Dayton II*).

The lower courts, but not yet the Supreme Court, have decided such issues as the right of the handicapped to public school education, e.g., *Fialkowski v. Shapp*, 405 F. Supp. 946 (E.D. Pa. 1975); *Mills v. Board of Educ.*, 348 F. Supp. 866 (D.D.C. 1972); *Pennsylvania Ass'n for Retarded Children v. Pennsylvania*, 343 F. Supp. 279 (E.D.

However, the significant victories that students won, culminating in *Goss v. Lopez*²⁶ and *Wood v. Strickland*²⁷ in 1975, appear to have ended. In the past three terms the Supreme Court, led by Justices Powell — a former school board president — and Rehnquist, has reemphasized the importance of judicial restraint and reiterated its support for the discretion of school personnel to make important decisions that affect students' lives. Finding that schools were not within the ambit of the eighth amendment's ban on cruel and unusual punishment, the Court in *Ingraham v. Wright*²⁸ refused to bar corporal punishment or to require a hearing prior to its imposition. Speaking for a divided Court, Justice Powell noted that the Court was reviewing a "legislative judgment, rooted in history and reaffirmed in the laws of many States, that corporal punishment serves important educational interests."²⁹ Although he admitted that school authorities could choose to abandon physical discipline and that its elimination might be welcomed as a societal advance, he upheld the right of school officials to administer it "reasonably." In *Board of Curators of the University of Missouri v. Horowitz*³⁰ the Court held that due process did not require a hearing prior to the dismissal of a medical student completing her fourth year of training, even though it found that "dismissal from a graduate medical school was more severe than the 10-day suspension to which the high school students were subjected in *Goss*."³¹ *Goss* was a

Pa. 1972); the proper identification of learning disabled children, *e.g.*, *Frederick L. v. Thomas*, 419 F. Supp. 960 (E.D. Pa. 1976), *aff'd*, 557 F.2d 373 (3d Cir. 1977), *returned to state court for assessment of damages*, 578 F.2d 513 (3d Cir. 1978); the right of schools to exclude handicapped children, *Stuart v. Nappi*, 443 F. Supp. 1235 (D. Conn. 1978); the distribution of sex questionnaires to pupils by high school newspaper staff, *Trachtman v. Anker*, 563 F.2d 512 (2d Cir. 1977), *cert. denied*, 435 U.S. 925 (1978); the liability of school officials for failure to educate children to a certain level, *Doe v. San Francisco Unified School Dist.*, Civ. No. 36851 (Cal. Ct. App., Aug. 6, 1976); the assessment of minority group children, *Lora v. Board of Educ.*, 456 F. Supp. 1211 (E.D.N.Y. 1978); *Larry P. v. Riles*, 343 F. Supp. 1306 (N.D. Cal. 1972) (order granting preliminary injunction), *aff'd*, 502 F.2d 963 (9th Cir. 1976), No. C-71-2270 R.F.P. (N.D. Cal. Oct. 16, 1979) (decision on merits), *appeal docketed*, No. 80-4027 (9th Cir. Jan. 17, 1980); *Hobson v. Hansen*, 269 F. Supp. 401 (D.D.C. 1967), *aff'd sub nom. Smuck v. Hobson*, 408 F.2d 175 (D.C. Cir. 1969).

26. 419 U.S. 565 (1975).

27. 420 U.S. 308 (1975).

28. 430 U.S. 651 (1977). See generally Comment, *Ingraham v. Wright: Corporal Punishment in Schools Passes Constitutional Tests*, 37 Md. L. REV. 594 (1978).

29. 430 U.S. at 681.

30. 435 U.S. 78 (1978).

31. *Id.* at 86 n.3.

disciplinary determination requiring fact-finding, while *Horowitz* was denominated an academic evaluation:

Such a judgment is by its nature more subjective and evaluative than the typical factual questions presented in the average disciplinary decision. . . . The determination whether to dismiss a student for academic reasons requires an expert evaluation of cumulative information and is not readily adapted to the procedural tools of judicial or administrative decisionmaking.

. . . We decline to further enlarge the judicial presence in the academic community and thereby risk deterioration of many beneficial aspects of the faculty-student relationship. We recognize, as did the Massachusetts Supreme Judicial Court over 60 years ago, that a hearing may be "useless or harmful in finding the truth as to scholarship."³²

Horowitz, then, may have some important implications for the use of psychological tests. Where tests are employed for academic purposes, as almost all are, courts may be reluctant to scrutinize their use. Further, *Horowitz* emphasizes the need for unfettered "expert evaluation of cumulative information." To a lesser degree, *Regents of the University of California v. Bakke*³³ also stands for judicial restraint. *Bakke* permits unencumbered individualized

32. *Id.* at 89-90 (citations omitted). *Cf. Regents of Univ. of Cal. v. Bakke*, 438 U.S. 265, 319 n.53 (1978) (opinion of Powell, J.) (racial quotas unconstitutional but racial considerations permitted in the admission of students to professional training; "So long as the university proceeds on an individualized, case-by-case basis, there is no warrant for judicial interference in the academic process."). One student commentator has similarly stated, "[M]ost courts have been lenient in dealing with classifications affecting educational rights; school boards have been given wide latitude so long as their policies remain untainted by racial motivation." *Judicial Supervision*, *supra* note 8, at 1512; see cases collected at *id.* n.13.

33. 438 U.S. 265 (1978). The *Bakke* opinion was severely fragmented. Justices Stevens, Rehnquist, Stewart, and Chief Justice Burger found that the Medical School of the University of California at Davis had violated the ban on racial discrimination in Title VI of the Civil Rights Act of 1964, 42 U.S.C. §§ 2000d to 2000d-4 (1976). *Id.* at 408-21 (Stevens, Stewart, Rehnquist, JJ., & Burger, C.J.). Justices Brennan, White, Marshall, and Blackmun disagreed, finding that Title VI's prohibition on discrimination was no broader than that of the equal protection clause of the fourteenth amendment, and that the Davis program was not unconstitutional. *Id.* at 328-79 (Brennan, White, Marshall & Blackmun, JJ.). In his pivotal opinion, Justice Powell agreed with the Brennan group that the question was a constitutional one, but concluded that the Davis program was unconstitutional. He found, however, that race *could* be treated as a factor in university admission decisions. *Id.* at 281-320 (Powell, J.). There was thus a 5-4 majority for each of two holdings: that Allan Bakke should be ordered admitted to the medical school, and that a school can constitutionally consider race in its admission procedure, at least under certain circumstances. See generally 92 HARV. L. REV. 131 (1978).

academic decisionmaking so long as impermissible racial or ethnic considerations are not part of the decision and the process itself violates no statute. It is just these purposes for which tests have been deemed useful — case-by-case appraisals of educational deficits and instructional needs performed by designated professionals based on assembled data.

By asserting the broader implications of the cited cases, a major constraint on the use of psychological tests — the restriction with regard to racial and ethnic impact — has also been stated.³⁴ As shall be shown, testing has been intimately involved in the history of racial and ethnic discrimination in the public schools.

TESTING AND DISCRIMINATION³⁵

The Pupil Placement Laws

The Supreme Court's ringing declarations in *Brown v. Board of Education*³⁶ began the slow process of desegregation in the South. Many southern states attempted to forestall the process by creating mechanisms that would prevent black children from attending previously all-white schools.³⁷ One such mechanism was the so-

34. *Educational Testing*, *supra* note 8, at 386-87:

To seek constructive judicial review of special education programs requires the court to step outside its usual realm of expertise. There is great danger in the temptation to deal with problems formalistically or mechanically through the routine application of legal doctrine The issue of testing will be an exceedingly difficult area for courts to consider, especially when the issue becomes one of the biases in the operation of special education programs. . . . [T]he courts cannot inquire into whether a given test is good or bad, but must determine whether a system is reasonably calculated to further the purposes of education. If a system is not furthering the purposes of special education, and if it is predominantly minority group children who are affected by the flawed system, then the court must achieve the delicate task of protecting interests without usurping the functions of the school board.

35. This section will be devoted almost exclusively to race and the attempted use of tests to hamper integration efforts in the South. During the early part of the twentieth century, some psychologists helped foster ethnic discrimination by arguing incorrectly that low test scores of certain immigrants were attributable to genetic differences rather than to socioeconomic class differences. See L. KAMIN, *supra* note 7; Kamin, *supra* note 7.

36. 347 U.S. 483 (1954). The Supreme Court declared that "segregation is a denial of . . . equal protection," *id.* at 495; that public education "must be made available to all on equal terms," *id.* at 493; and that to separate black children from white "solely because of their race generates a feeling of inferiority . . . that it may affect their hearts and minds in a way unlikely ever to be undone," *id.* at 494.

37. "The myriad of new pupil assignment laws, 'interposition' plans, and other ingenious schemes demonstrated the truth of the popular saying — 'as long as we can legislate we can segregate.'" Yudof, *supra* note 8, at 465. See generally Meador, *The Constitution and the Assignment of Pupils to Public Schools*, 45 VA. L. REV. 517 (1959); Note, *The Federal Courts and Integration of Southern Schools: Troubled Status of the Pupil Placement Acts*, 62 COLUM. L. REV. 1448 (1962).

called pupil placement acts,³⁸ which relied heavily on the use of intelligence and achievement tests.³⁹

The pupil placement acts prescribed criteria by which local school boards could make decisions about students seeking intradistrict school transfers. In reality, they served to screen black applicants and block their transfers. Among such apparently innocuous criteria as the availability of space and transportation, also considered were scholastic aptitude, intellectual ability, and the effect of admission upon prevailing academic standards at the prospective school.⁴⁰ When Alabama's School Placement Act was held constitutional,⁴¹ school systems received the imprimatur of the Supreme Court to consider intelligence and achievement when evaluating transfer applications of black students.⁴²

38. ALA. CODE tit. 52, §§ 61(1)-(12) (1958); ARK. STAT. ANN. §§ 80-1519 to -1547 (1960); FLA. STAT. ANN. § 230.232 (1960); LA. REV. STAT. §§ 17:101-110 (Supp. 1961); MISS. CODE ANN. §§ 6334-01 to -08 (Supp. 1960); N.C. GEN. STAT. §§ 115-176 to -179 (1960); S.C. CODE §§ 21-230, 21-247 to -247.7 (Supp. 1960); TENN. CODE ANN. §§ 49-1741 to -1764 (Supp. 1962); TEX. REV. CIV. STAT. art. 290(a) (Supp. 1962); VA. CODE ANN. §§ 22-232.1 -232.31 (Supp. 1962).

39. The use of tests to prevent integration did not begin, however, in the South. In 1869 Indiana provided separate schools for black and white children, but the schools for black children did not reach the upper grades. Indiana law permitted black children to enter the advanced grades of all-white schools only if they could pass an examination. When a black child contested the process that denied him admission to the graded white school, the Indiana Supreme Court concluded that since "no such examination was passed . . . this bars a right to admission . . . on the ground of proficiency in study." *State ex rel. Mitchell v. Gray*, 93 Ind. 303, 305 (1883). The court apparently considered itself powerless to act unless corruption or malice were shown: "The Legislature has ample power to regulate these matters, and it is for us to respect and enforce the law as it is written on our statute books." *Id.* "[C]ourts cannot undertake the government of . . . school affairs." *Id.* at 304.

40. *Dove v. Parham*, 181 F. Supp. 504, 510 n.5 (E.D. Ark. 1960). Other criteria included the effect of admission of the child upon the academic progress of other students, the possibility of breaches of the peace within the community, the child's home environment, and the morals, conduct, health and personal standards of the pupil. *Id.*

41. *Shuttlesworth v. Birmingham Bd. of Educ.*, 162 F. Supp. 372 (N.D. Ala.), *aff'd*, 358 U.S. 101 (1958) (three judge district court asserted that the law "furnishes the legal machinery for an orderly administration of the public schools in a constitutional manner by the admission of qualified pupils upon a basis of individual merit without regard to their race or color."). The Supreme Court made clear that it was affirming the judgment only "upon the limited grounds on which the District Court rested its decision," i.e., that in advance of its application it was not possible to say that the law, constitutional on its face, would be unconstitutionally administered. 358 U.S. at 101.

42. Though the pupil placement acts were the first formal state-initiated means for delaying integration, less formal local policies had already been reviewed by the courts. The Dallas, Texas Board of Education had found that there were differences in the scholastic aptitudes of white and black children in elementary and secondary grades. On that basis it refused to order the end of segregated schools. In *Borders v.*

The acts were sustained as expressions of state sovereignty by which states could experiment with methods of achieving gradual and well-regulated integration.⁴³ Intellectual criteria were judged to be "unquestionably valid in and of themselves,"⁴⁴ and the use of tests was uniformly upheld.⁴⁵ The plaintiffs never challenged the validity of the tests, nor did the courts. The only concern of the federal courts at the time was whether the placement acts were used to make transfer decisions "purely" upon racial considerations.⁴⁶ Test use was enjoined only when the courts found that the tests were administered solely to black children who wanted to attend "white" schools and thus were employed as overt attempts to avoid desegregation.⁴⁷ Occasionally, a court lost patience as it was forced

Rippy, 247 F.2d 268 (5th Cir. 1957), the Fifth Circuit found that black children were excluded from schools of their choice solely on the basis of race and thus were denied equal protection and due process. It stated, however, that "pupils may, of course, be separated according to their degree of advancement or retardation, their ability to learn, . . . or for any other legitimate reason, but each child is entitled to be treated as an individual without regard to his race or color." *Id.* at 271. *Accord*, Youngblood v. Board of Pub. Instr'n, 230 F. Supp. 74, 76 (N.D. Fla. 1964).

43. Parham v. Dove, 271 F.2d 132 (8th Cir. 1959), *remanded*, 181 F. Supp. 504 (E.D. Ark. 1960).

44. Dove v. Parham, 181 F. Supp. 504, 520 (E.D. Ark. 1960).

45. *E.g.*, Jones v. School Bd., 278 F.2d 72, 75 (4th Cir. 1960) ("It is not contended by the appellants . . . that intelligence or scholarship attainments may never be properly applied in determining the particular schools that children shall attend. . . . In the absence of a showing that these factors are used in such a way as to deprive individuals of their constitutional rights, they are, of course, not objectionable on constitutional grounds."); Evans v. Ennis, 281 F.2d 385, 395 (3d Cir. 1960) ("It was not our intention . . . to exempt the . . . plaintiffs who may presently actively seek integration from the usual processing of the school system relating to their capabilities, scholastic attainments"); *see* Norwood v. Tucker, 287 F.2d 798, 803 (8th Cir. 1961). Among the tests ratified were the California Achievement Test, Pettit v. Board of Educ., 184 F. Supp. 452 (D. Md. 1960); the California Test of Mental Maturity, Dodson v. School Bd., 289 F.2d 439 (4th Cir. 1961); the California Test of Personality, Augustus v. Board of Pub. Instr'n, 306 F.2d 862 (5th Cir. 1962); and the Iowa Silent Reading Test, Dodson v. School Bd., 289 F.2d 439 (4th Cir. 1961).

46. Parham v. Dove, 271 F.2d 132, 137 (8th Cir. 1959).

47. *E.g.*, Norwood v. Tucker, 287 F.2d 798, 808-09 (8th Cir. 1961) (where blacks were compelled to undergo a series of tests covering education, psychology, religion, ethics, and law the court found that the criteria for transfer were being applied "for the purpose of impeding, thwarting, and frustrating integration," and held that "the standards and criteria of the pupil assignment law must be applied objectively in the making of initial assignments of all students . . ."); Dodson v. School Bd., 289 F.2d 439, 443 (4th Cir. 1961) (academic criteria applied only to black high school pupils is "patently discriminatory"); Jones v. School Bd., 278 F.2d 72, 77 (4th Cir. 1960) ("If the criteria should be applied only to Negroes seeking transfer or enrollment in particular schools and not to white children, then the use of the criteria could not be sustained."); *see* Augustus v. Board of Pub. Instr'n, 306 F.2d 862, 869 (5th Cir. 1962); Youngblood v. Board of Pub. Instr'n, 230 F. Supp. 74, 76 (N.D. Fla. 1964); Dove v. Parham, 181 F. Supp. 504, 520 (E.D. Ark. 1960). *Cf.* Calhoun v. Latimer, 321 F.2d 302, 316-17 (5th Cir.

to review one scheme after another that some inventive school board had created:

[T]he District has . . . made its processes of application of the statute consist in having applicants for transfer subjected to such devices as the California Mental Maturity Test, the Iowa Silent Reading Test, the Otis Quick Scoring Test of Mental Ability, the California Language Test, the Bell Adjustment Inventory, and other such things — which, at least in the elementary area of public education, are new adornments upon the entrance doors to school houses and classrooms.

. . . [I]n what the District has done and proposes to continue doing, application of these devices is not going to be made to the students generally of the system but only to such individuals as undertake to engage in application for a transfer — which in the realities of the District here simply means, to Negro students seeking to enter a white school.⁴⁸

But, in the main, courts acted gingerly in these early cases. As stated by one commentator, “[M]ost federal courts were careful to delve into matters of school policy only after drawing the parameters of their inquiry to avoid any encroachment upon the school boards’ discretionary authority.”⁴⁹ They made clear that their function was not “to make educational policy, [to] resolve conflicts in educational theory,”⁵⁰ or to declare tests “unfair or unacceptable as a matter of law.”⁵¹ Only when “educational principles and theories serve[d] to justify [the preservation of] an existing system of imposed segregation”⁵² were testing programs enjoined as barriers to the vindication of the constitutional rights secured in *Brown*. Test administration that applied to everyone and that led to regrouping among and

1963) (“Personality interviews to determine probable success or failure in the schools to which transfer and assignment is sought may not be utilized where such a practice relates only to Negro pupils . . . nor may any scholastic requirement whatever be used where applied only to Negro students seeking transfer.”), *vacated and remanded on other grounds*, 377 U.S. 263 (1964). The Atlanta School Board used another dissembling device. While achievement tests were given routinely to all children, those requesting transfers had to attain a grade level score at least equal to the average of the class in the school to which transfer was requested. This practice was also declared unlawful in *Calhoun*. *Id.* at 304-05.

48. *Dove v. Parham*, 282 F.2d 256, 260 (8th Cir. 1960).

49. *Cultural Bias*, *supra* note 8, at 1036.

50. *Dove v. Parham*, 181 F. Supp. 504, 519 (E.D. Ark. 1960).

51. *Jones v. School Bd.*, 179 F. Supp. 280 (E.D. Va. 1959).

52. *Dove v. Parham*, 282 F.2d 256, 258 (8th Cir. 1960).

within schools was free from judicial scrutiny.⁵³ In the early 1960's, when the courts were attempting to make school systems begin the process of desegregation, charges that tests themselves were racist, culturally biased, and discriminatory had not yet been heard.

The Ghost of Brown

In what has become a famous footnote, the Supreme Court in *Brown v. Board of Education*⁵⁴ cited the works of several behavioral scientists to support its conclusion that segregation retarded the educational and cognitive development of black children.⁵⁵ The reference to the studies created a controversy that still exists concerning their relevance and validity and the methodology that produced them.⁵⁶

The first and most significant test of *Brown's* footnote eleven took place in Georgia when black children sought to enjoin a county board of education's continued operation of a de jure dual school system. In *Stell v. Savannah-Chatham County Board of Education*,⁵⁷ two white children intervened on behalf of white students as a class to show that the alleged segregation was not determined solely on the basis of color "but rather upon racial traits of educational significance as to which racial identity was only a convenient index."⁵⁸ The white children sought to prove that differences in learning rates, cognitive ability, behavioral traits, and capacity for education in general were so great that not only was it impossible for black and white children to be educated effectively in the same room but that "to congregate children for such diverse traits in schools . . . would seriously impair the educational opportunities of both white and Negro and cause them grave psychological harm."⁵⁹

53. See, e.g., *id.* at 260. It may not be a coincidence that "from 1957 until 1966, the use of standardized tests doubled." *Equal Protection*, *supra* note 8, at 900 (citing W. MEHRENS & J. LEHMAN, *STANDARDIZED TESTS IN EDUCATION* 5 (1969)).

54. 347 U.S. 483 (1954).

55. *Id.* at 494 n.11.

56. E.g., Cahn, *Jurisprudence*, 30 N.Y.U. L. REV. 150 (1955); Levin, *School Desegregation Remedies and the Role of Social Science Research*, 42 L. & CONTEMP. PROB. 1 (1978); Levin & Moise, *School Desegregation Litigation in the Seventies and the Use of Social Science Evidence: An Annotated Guide*, 39 L. & CONTEMP. PROB. 50 (1975); Yudof, *supra* note 8, at 437-39, 444-45. See generally R. KLUGER, *SIMPLE JUSTICE* 705-07 (1975).

57. 220 F. Supp. 667 (S.D. Ga. 1963), *rev'd*, 333 F.2d 55 (5th Cir.), *cert. denied*, 379 U.S. 933 (1964).

58. *Id.* at 668.

59. *Id.*

To prove their contentions the intervenors called several expert witnesses, among whom were two professors of psychology, R. Travis Osborne and Henry Garrett. Osborne had evaluated all black and white students in the sixth, eighth, tenth, and twelfth grades in the defendant school district over approximately a ten-year period using a battery consisting of the California Achievement Tests (CAT) and the California Mental Maturity Tests (CTMM).⁶⁰ On the basis of his assessment, Osborne testified that black children were two to four years behind their white peers in reading and arithmetic and had a median I.Q. of 81, compared to a median I.Q. of 103 for white children, despite the fact that black teachers had more education and experience than did their white counterparts. Further, Osborne stated that black children fell behind a white control group matched for age and mental ability in school achievement. He concluded from these data that educational separation for the two groups was necessary if, as the district court paraphrased it, "the schools are to endeavor to adapt to the different learning potentials of each."⁶¹

Garrett verified Osborne's results and stated they were representative of the nation as a whole. He asserted that to educate both races together would be "a source of frustration [that would] be compensated for by anti-social class behavior."⁶² The court summarized Garrett's conclusions as follows:

[D]ifferences in educability between Negro and white children were inherent and . . . only minor changes could be achieved by educational readjustment or other environmental change. There was no scientific possibility that learning rate differences of the degree shown by Dr. Osborne's tests and the confirming national studies were either caused by or could be substantially altered by the students' environment.⁶³

60. Osborne characterized the CAT as a set of nationally accepted achievement tests in reading comprehension and vocabulary, mathematical reasoning and fundamentals, and the application of mathematical concepts. *Id.* The CTMM was described as a "nationally accepted standard indicator of the ratio between mental and chronological age. . . ." *Id.* However, a recent critique of the CAT asserts that it does not measure what it purports to measure. J. SALVIA & J. YSELDYKE, *supra* note 4, at 133. For generally lukewarm reviews of the CTMM, see L. CRONBACH, *supra* note 7, at 277; G. ROBB, L. BERNARDONI & R. JOHNSON, *supra* note 7, at 289-91.

61. 220 F. Supp. at 671.

62. *Id.* at 672.

63. *Id.* It may be some indication of the naivete of the plaintiffs' attorneys who litigated these early desegregation-testing cases that they stipulated that "the opinions of Dr. Garrett were authoritative in his field of experimental and differential psychology." *Id.*

The district court admitted this and other testimony⁶⁴ because it concluded that the findings in *Brown* with regard to psychic damage suffered by black children did not control in the case before it. It held that "the existence or non-existence of injury to white or black children from integrated or segregated schooling is a matter of fact for judicial inquiry and was so treated in *Brown*."⁶⁵ Thus, because under the principle of *stare decisis* future courts were bound only with regard to matters of law, the issue of injury was justiciable.⁶⁶

The court concluded: "[T]he differences in test results between the white and Negro students is [*sic*] attributable in large part to hereditary factors [and] failure to attain the existing white standards would create serious psychological problems of frustration on the part of the Negro child, which would require compensation by attention-creating anti-social behavior."⁶⁷ Although the court acknowledged that one-third of the black children had achievement and intelligence test scores comparable to whites, it found that "selective integration would cause even greater psychological harm to the individual Negro children [who] would be inescapably conscious of total social rejection by the dominant group."⁶⁸ As for the lesser achieving blacks, integration would "greatly increase any existing sense of inferiority [causing] substantial and irremovable psychological injury"⁶⁹ Believing that the white intervenors

64. Another psychologist testified that the rate of truancy and antisocial behavior generally could be diminished if a classroom were organized "around a socially homogeneous group of children having relatively similar learning rates." This opinion was based on data that showed that "one-third of all Negro truants gave inability to keep abreast of their school work as the reason for their running away from home." *Id.* at 675. Other witnesses claimed that variations in intellectual ability between blacks and whites were innate because of differences in brain size, proportion, and structure, and that racial mixing increased pre-existing hostility especially where there were differences in physical appearance and learning rates. *Id.* at 673-74.

65. *Id.* at 678.

66. The district court criticized the conclusions reached in *Brown*. *Id.* at 678-80. As noted, text accompanying note 56 *supra*, there has been controversy since 1954 as to the need for and meaning of the psychological data. One commentator has concluded that finding an empirical basis for the decision was unnecessary. Yudof, *supra* note 8, at 437. "The evidence is too complex, unreliable, unpersuasive, and contradictory to allow the fashioning of broad constitutional principles, even if it were possible to disentangle the psychological and educational components of segregation and attribute injurious consequences to either." *Id.* at 445. "*Brown* and its progeny are fundamentally based on ethical principles . . . [i.e.,] normative, or mixed normative and factual, statements about how a just society should function." *Id.* at 446.

67. 220 F. Supp. at 683.

68. *Id.* at 684.

69. *Id.*

had successfully disproved the factual findings in *Brown*, the court dismissed the black children's complaint, concluding that "classification . . . on the basis of coherent groups having distinguishable educability capabilities" was reasonable and consonant with the equal protection principle that reasonable classifications are those that secure "the maximum result in the educational process for all students and the minimum injury to any."⁷⁰

The black plaintiffs in *Stell* appealed, and the Fifth Circuit reversed.⁷¹ Judge Griffin Bell, speaking for the court, found that *Brown* was not limited to its facts and that all courts were bound by its holding that segregated schools were constitutionally impermissible. He admonished the judiciary "that no federal court may refrain from acting as required by that decision even if such court should conclude that the Supreme Court erred either as to its facts or as to the law."⁷² Nevertheless, he deferred to the judgment of educators: "[T]here is no constitutional prohibition against an assignment of individual students to particular schools on a basis of intelligence, achievement or other aptitudes upon a uniformly administered program, but race must not be a factor in making the assignments. However, this is a question for educators and not courts."⁷³ Savannah's plan was faulty not because of the testing program or the inadequate bases on which the experts drew their conclusions, but because black children were to be classified regardless of their individual abilities. "Therein is the discrimination. The individual Negro student is not to be treated as an individual and allowed to proceed along with other individuals on the basis of ability alone without regard to race."⁷⁴

70. *Id.* The court did not examine the purposes of grouping to discern whether they were merely subterfuges for de jure racial segregation. Under equal protection doctrine racial classifications are given strict scrutiny and upheld only if the state can show that there is a compelling justification for them and that there is no less drastic means for accomplishing the state's asserted goal. *E.g.*, *McLaughlin v. Florida*, 379 U.S. 618 (1964). Most other classifications that differentiate among groups of persons are subjected only to the more relaxed level of the rational basis test, the one the court invoked here. See note 185 *infra*.

71. 333 F.2d 55 (5th Cir. 1964).

72. *Id.* at 61-62.

73. *Id.*

74. *Id.* at 62. While the *Stell* case was pending in the Fifth Circuit, a federal district court in northern Florida entertained a similar desegregation case. *Youngblood v. Board of Pub. Instr'n*, 230 F. Supp. 74 (N.D. Fla. 1964). Opposing a motion for summary judgment by the black plaintiffs, the school district — relying on the district court's opinion in *Stell* as precedent — proffered what the court called "a considerable volume of statistics . . . tending to establish . . . that there are inherent racial differences in intelligence, aptitude, and in rate of intellectual attainment . . . justify[ing] the use of a racial index as one criteria in making assignment of pupils to

Just three weeks after the Fifth Circuit's ruling that data on the injurious effects of integration were immaterial, a district court in Mississippi permitted the defendant school district and white intervenors in a suit very similar to *Stell* to present evidence offered by "seven distinguished scientists" as to the physical, intellectual, and educational inferiority of black children.⁷⁵ Apparently more sophisticated in court than its predecessors, this school system sought to counter the argument that the I.Q. tests used to justify the separation of the races were culturally biased:

A special test was . . . made to determine whether intelligence tests unduly favored white pupils because of containing cultural questions which might be less familiar to Negro families. The results were . . . contrary to the "Cultural Hypothesis" in that the Negro group scored relatively higher on those questions which had been rated by educators as being highly cultural in content.⁷⁶

On the basis of what it called "overwhelming, undisputed, and unchallenged evidence," the court concluded that the "differences between Caucasians and Negroes are genetically determined and cannot be changed materially by the environment."⁷⁷ Yet, despite these findings and its belief that the determinations in *Brown* were matters of fact that could be relitigated, the court reluctantly found itself bound by the Fifth Circuit's holding in *Stell* and permanently enjoined the operation of the dual school system. The school district

specific schools." *Id.* at 76. Predicting the Fifth Circuit's reversal, the court concluded that "*Stell* . . . is not in accord with the requirements of the . . . Fifth Circuit's prior opinions . . .," *id.* at 75, and ignored the data. It granted the plaintiff's summary judgment motion as to the pupil assignment plan, stating that any procedure that embodied "a universal testing basis for assignment may not be administered in a manner which would defeat the essential requirement that factors of race are not to be considered." *Id.* at 76.

75. *Evers v. Jackson Mun. Separate School Dist.*, 232 F. Supp. 241, 246 (S.D. Miss. 1964), *aff'd*, 357 F.2d 653 (5th Cir.), *cert. denied*, 384 U.S. 961 (1966). The scientists included several psychologists, who were not named in the court's opinion. The evidence was very similar to that presented in *Stell*: methodological criticism of the psychological studies cited in *Brown*, proof of increased hostility between the races as the result of compulsory intermixing in groups having a high degree of self-identity, and disparity in achievement and ability. The court also heard testimony concerning a longitudinal study of black and white children initially matched for I.Q. in which it was found after three years that the blacks were "one year apart in terms of I.Q., measured by the same tests." *Id.* at 247.

76. *Id.* at 248. For further discussion of test bias, see text accompanying notes 111 to 148 *infra*.

77. *Id.* at 251.

appealed to the Fifth Circuit, but one of that circuit's panels, irritated by cases that "tax the patience of the court," tersely affirmed the lower court's opinion, interring "the popular myth that Brown was decided for sociological reasons untested in a trial" ⁷⁸

*Hobson v. Hansen*⁷⁹ and the Attack on Group Ability Tests

Almost without exception, the tests and results that led expert witnesses to conclude that black children were genetically inferior went unchallenged by attorneys fighting to enforce desegregation. They argued only that such evidence and the conclusions drawn from it were immaterial and irrelevant,⁸⁰ and appellate courts left such issues unexamined. States, now educated by decisions in the first decade after *Brown*, repealed or drastically altered the pupil placement acts, and school systems cautiously tested all children and developed "freedom of choice" plans, which, on the surface, made it easier for blacks to attend desired schools. As long as race was not the evident criterion on which decisions were made, judicial restraint and support for school system discretion continued.

The inevitable conclusion from these cases was that testing to measure academic ability was perfectly permissible. For example, a federal court in the South in 1966 stated: "[T]he practice of separating study groups or classes into accelerated or slow sections is a matter for educators. 'Discrimination' bears no stigma in connotation. It is 'racial discrimination' in public education which must be excised from the American culture."⁸¹ One district court apparently read prior opinions as not only permitting but prescrib-

78. *Jackson Mun. Separate School Dist. v. Evers*, 357 F.2d 653, 654 (5th Cir.), *cert. denied*, 384 U.S. 961 (1966). See *United States v. Board of Educ.*, 301 F. Supp. 1024 (S.D. Ga. 1969). There the district court permitted the testimony of psychologists concerning the disparity between black and white children with regard to intelligence and achievement (on the CAT and CTMM) and the conclusion that cultural difference and socioeconomic status have little to do with the disparity. The court reluctantly ignored the data in arriving at its decision to enjoin the further operation of segregated schools in Lincoln County, finding itself bound by *Stell* and *Jackson*. *Id.* at 1029. However, the court reacted angrily to those constraints: "Under latter-day Fourteenth Amendment interpretation, scholastic aptitude means nothing. Total integration of schools, regardless of consequences to the system, is all that counts. The higher courts apparently look no further." *Id.*

79. 269 F. Supp. 401 (D.D.C. 1967).

80. *E.g.*, *Evers v. Jackson Mun. Separate School Dist.*, 232 F. Supp. 241, 244 (S.D. Miss. 1964).

81. *Miller v. School Dist. No. 2, Clarendon County, S.C.*, 256 F. Supp. 370, 375 (D.S.C. 1966).

ing testing.⁸² It approved a portion of a school board's integration plan prohibiting students from being assigned or transferred to any school or class in which the mean I.Q. score exceeded that of the student. To effectuate the plan the court "enjoined and required [the school system] to maintain and enforce distinctions based upon age, mental qualifications, intelligence, achievement and other aptitudes upon a uniformly administered program."⁸³ The injunction was quickly dissolved, however, when the Fifth Circuit reversed, stating that "the order of the trial court directing the Board of Education to assign pupils according to intelligence tests was beyond the power of the court."⁸⁴

It was against this background that Judge Skelly Wright's revolutionary opinion in *Hobson v. Hansen* appeared.⁸⁵ The case dealt with the constitutionality of intraschool district disparities in financial and educational resources that favored white children.⁸⁶ Its most important issue, however, was the propriety of placement through standardized tests of a disproportionate number of black children in lower academic tracks and white children in upper tracks. The defendant, the District of Columbia school system, cited *Stell* and its progeny as support for the validity of its procedures.⁸⁷ The court distinguished those cases, calling them "completely inapposite . . . because in none of them did the courts have occasion

82. *Stell v. Savannah-Chatham County Bd. of Educ.*, 255 F. Supp. 83 (S.D. Ga. 1965), *on remand from* 333 F.2d 55 (5th Cir. 1964), *rev'g* 220 F. Supp. 667 (S.D. Ga. 1963), *discussed at text* accompanying notes 37 to 74 *supra*.

83. *Id.* at 99.

84. *Stell v. Board of Pub. Educ.*, 387 F.2d 486, 491 (5th Cir. 1967). The court added cryptically: "No person has the constitutional right to attend schools in classes which are divided according to ability or intelligence quotients." *Id.* It went on to admonish the lower courts that "there is no constitutional authority for federal courts to impose their ideas of school management on boards of education as to these matters." *Id.* But, without realizing future implications of its action, the Fifth Circuit permitted the school to assign children to special classes or schools for the physically handicapped, retarded, and gifted on bases related to the function of those schools. *Id.* at 496.

85. 269 F. Supp. 401 (D.D.C. 1967), *aff'd sub nom.* *Smuck v. Hobson*, 408 F.2d 175 (D.C. Cir. 1969). Commentators have been uniform in their judgment of the case's importance. "[T]he most far-reaching judicial foray into school classification . . ." Sorgen, *supra* note 8, at 1139; "[T]he . . . Court . . . reached a conclusion so profound and far-reaching that, if followed by other courts, . . . could radically alter the educational system of almost every urban school in the nation." *Ability Grouping*, *supra* note 8, at 150; "[M]ay presage a period of increased judicial scrutiny of educational policies. . . ." *Judicial Supervision*, *supra* note 8, at 1511.

86. See Kirp, *Law, Politics, and Equal Educational Opportunity*, *supra* note 8, at 129; *Legal Implications*, *supra* note 8, at 737.

87. 269 F. Supp. at 514 n.212. See text accompanying notes 57 to 74 *supra*.

to consider whether the tests used were in fact accurate in ascertaining innate ability."⁸⁸

The track system in Washington, D.C. was instituted two years after the Supreme Court decided *Bolling v. Sharpe*,⁸⁹ a companion case to *Brown*.⁹⁰ Unlike many southern school systems, the District had integrated its schools very quickly.⁹¹ Soon after the integration was implemented, however, teachers began reporting severe academic deficiencies in black children.⁹² It was an attempt to remedy these deficiencies that led the superintendent of schools to institute the track system. A commentator has described the system:

The idea was to segregate children of the same ability in the same classes so that all could attain optimum development. Each student was placed in Honors, Regular, General, or Basic^[93] program with emphasis on college preparation in the higher groups and on blue collar jobs in the lower. There was to be some overlap between tracks, e.g., a student might be in Honors mathematics and only Regular reading. Also there was to be continuous re-evaluation of the students' ability with correction of any remediable education problems accounting for present poor performance.⁹⁴

Despite its recognition of the superintendent's good intentions,⁹⁵ what led to the court's condemnation of ability grouping was its finding of disproportionality. "As a general rule, in those schools with a significant number of white and Negro students a higher proportion of the Negroes will go into the Special Academic

88. 269 F. Supp. at 514 n.212.

89. 347 U.S. 497 (1954).

90. The cases in *Brown* arose in four states, *Bolling* in the District of Columbia. Although the District was not bound by the fourteenth amendment's prohibition of denial of equal protection through state action, in *Bolling* the Supreme Court interpreted the fifth amendment's prohibition of deprivation of due process by the federal government so as to encompass the principle of equal protection. 347 U.S. at 498-99.

91. By September 1954, 73% of the schools were racially mixed. *Hobson v. Hansen*, 269 F. Supp. 401, 442 (D.D.C. 1967).

92. The court called them "the casualties of *de jure* segregation." *Id.* at 443.

93. "Basic" was also called the Special Academic Track, a euphemism for classes for the educable mentally retarded. The criteria for placement in the Special Academic Track were inability to keep up with the normal curriculum, emotionally disturbed behavior, an I.Q. of 75 or below, and substandard performance on achievement tests. *Id.* at 446-48.

94. *Legal Implications*, *supra* note 8, at 738.

95. The court found that the superintendent was motivated primarily by a desire to correct an educational problem and that the track system was not merely a subterfuge to avoid desegregation. It stated that it was not resting "its decision on a finding of intended discrimination." 269 F. Supp. at 443. *See also id.* at 512 n.208.

[Educably Mentally Retarded (EMR)] Track than will white students."⁹⁶

What concerned Judge Wright, and what is pertinent here, was his finding that "the proper operation of the track system practically demands reliance on test scores."⁹⁷

[B]etween kindergarten and 12th grade a student will receive the following group tests: (1) a first-grade readiness test . . . ; (2) a reading and spelling achievement test (second grade); (3) one achievement test and one aptitude test in each of three grades (sixth, ninth, and 11th). In addition, optional aptitude tests may be given in the seventh, ninth, 10th or 12th grades; . . . Individual tests . . . are confined almost entirely to those students recommended for placement in the Special Academic Track.⁹⁸

Although the stated criteria for entrance into any one of the tracks included teacher and counselor evaluation of maturity, stability, physical condition, and grades, the court found that "testing looms as a most important consideration in making track assignments."⁹⁹ The findings that the track system had a negative impact on black children and that their placement was determined primarily by test scores triggered the court's extensive inquiry into

96. *Id.* at 456. The data on which the court's conclusion was based are reproduced below:

	School Year	Total School Enrollment		Ratio of Negroes to Whites in Special Academic Track	
		% Negro	% White	% Negro	% White
Elementary	1964	89.5	10.5	95.0	5.0
	1965	91.0	9.0	95.0	5.0
Junior High	1964	87.6	12.4	94.7	5.3
	1965	89.5	10.5	96.4	3.6

Id. (Table H). "Stated otherwise, as of October 1965, 3.3% of the white students in junior high school were in the Special Academic Track whereas 10.2% of the Negroes were in that Track." *Id.* at n.68. Based on frequency data found in Table A of the opinion, *id.* at 450, this author has calculated chi squares from the data in Table H. All are significant beyond the .01 level.

97. *Id.* at 475.

98. *Id.* at 476. Among the tests given were the Metropolitan Readiness and Achievement Tests, the Sequential Tests of Educational Progress, the School & College Ability Tests, the Stanford Achievement Tests, the Tests of General Ability, the Tests of Educational Ability, the Flanagan Aptitude Classification Tests, and the Otis Quick-Scoring Mental Ability Test, Beta. *Id.* at 518-19.

99. *Id.* at 475.

the nature and limitations of standardized tests.¹⁰⁰ As one commentator noted:

Analysis of the testing issue focused upon two concerns. First, the court examined test structure and found on the grounds of expert testimony that the skills measured by the test instruments were not innately intellectual but rather acquired through cultural experience. Accordingly, it stated, in order to demonstrate "ability" to learn the student must have had an opportunity to learn the requisite skills, which were in fact measures of the pupil's "cumulative experiences in his home, his community and his school." Since culture bias were built into the tests by standardization on norming groups which were predominantly white and middle-class, the disadvantaged child obviously would not score as well on the intelligence tests as his white counterpart from a background more conducive to the acquisition and development of the verbal and nonverbal skills measured by the test. . . .

The court next considered the impact of culture-biased testing upon the disadvantaged child. Elaborating upon the environmental and psychological indices of the disadvantaged child's situation within society, it explored possible educational effects of the testing upon students. In this regard, the court expressed concern over teacher evaluations based upon false opinions of intelligence projected by low test scores, a particular danger where teachers' universal reliance upon such tests was emphasized by a school policy mandating use of their results.¹⁰¹ The court found that the effects of mislabeling were cyclical, beginning with the lower evaluation of a child's ability and constant reinforcement of that evaluation.¹⁰²

After ninety pages of discussion, summary, and fact-finding, Judge Wright turned to the key legal issue of the case, whether ability grouping in the District of Columbia violated the equal protection clause. In *Carrington v. Rash*,¹⁰³ a case decided two years prior to *Hobson*, the Supreme Court had held that "mere classifica-

100. Another important finding was that despite the school's assertions to the contrary, the children in the lower tracks were almost always permanently relegated to those positions. *Id.* at 458-68. "Movement between tracks borders on the nonexistent." *Id.* at 463. See also *id.* at 512-13.

101. Another explanation for the low test scores offered by one of the experts and given credence by the court was test anxiety, *i.e.*, apprehension about ability to score well and fear of what the performance might mean to the test interpreters. The court found that such reactions were more likely to be found in black disadvantaged children than in their white advantaged counterparts. *Id.* at 482.

102. *Cultural Bias*, *supra* note 8, at 1044-45.

103. 380 U.S. 89 (1965).

tion . . . does not of itself deprive a group of equal protection.”¹⁰⁴ Following that decision, Judge Wright acknowledged that “not all classifications resulting in disparity are unconstitutional. If classification is reasonably related to the purposes of the government activity involved and is rationally carried out, the fact that persons are thereby treated differently does not necessarily offend.”¹⁰⁵ He then undertook to determine whether ability grouping, a form of classificatory activity, had a rational basis. At the outset, he decided that discrimination on the basis of ability could be defended only if judgments about ability were based on measures that assessed children’s capacity to learn, *i.e.*, their innate endowment, not their present skill levels. On that basis, ability grouping could “be reasonably related to the purposes of public education.”¹⁰⁶ But because the law “has a special concern for minority groups,” the existence of any practices that have the appearance of “invidious discrimination” would impose a “weighty burden” on the school system to explain why poor and black children disproportionately populate the lower tracks.¹⁰⁷ The only explanation the court would accept for the racial disparities was that the assessment devices on which the classifications were made accurately reflected students’ innate capacities to learn. It having already been found that tests did not do so, ability grouping was doomed:

While government may classify persons and thereby effect disparities in treatment, those included within or excluded from the respective classes should be those for whom the inclusion or exclusion is appropriate; otherwise the classification risks become wholly irrational and thus unconstitutionally discriminatory. It is in this regard that the track system is fatally defective, because for many students placement is based on traits other than those on which the classification purports to be based.¹⁰⁸

Judge Wright’s next words were to have a profound effect on the use of psychological tests during the next decade:

The evidence shows that the method by which track assignments are made depends essentially on standardized

104. *Id.* at 92.

105. 269 F. Supp. at 511.

106. *Id.* at 512.

107. *Id.* at 513. For a discussion of the level of scrutiny under equal protection analysis used by the court and the problem of burden of proof, see note 185 and text accompanying notes 186 to 206 *infra*.

108. *Id.*

aptitude tests which, although given on a system-wide basis, are completely inappropriate for use with a large segment of the student body. Because these tests are standardized primarily on and are relevant to a white middle class group of students, they produce inaccurate and misleading test scores when given to lower class and Negro students. . . . [T]hese students are in reality being classified . . . [on] factors which have nothing to do with innate ability.¹⁰⁹

Thus, Judge Wright ordered that "the track system . . . simply must be abolished"¹¹⁰ and permanently enjoined its operation.

A Critique of the Psychometric Underpinnings of Hobson

With one blow Judge Wright's decision in *Hobson* severely wounded two sacred cows, ability grouping and standardized testing. Although extensive comment on the educational implications of the former is not within the scope of this Article,¹¹¹ some brief remarks are in order. It is unclear whether Judge Wright ordered the dismantling of the track system under all conditions or only as it was then being implemented by the defendants. Although his order appears to call for abolition, Judge Wright had conceded earlier that ability grouping could be a rational practice. What the court complained of was its inflexibility, its stigmatizing effect on the black and the poor, its failure to provide students in lower tracks resources expended on those in upper tracks, and its failure to provide compensatory education for those in the General and Special tracks, thus relegating those students to permanent inequality. On appeal, Judge Wright's order was interpreted in the light most favorable to the board of education so that it would not hamper the flexibility of the new board or frustrate its attempts to improve the District's school system.¹¹² Thus, although the appellate court did not modify Judge Wright's order to abolish tracking, it did interpret

109. *Id.* at 514.

110. *Id.* at 515.

111. For educational and legal perspectives on ability grouping, see W. FINDLEY & M. BRYAN, *ABILITY GROUPING: 1970* (1971); M. GOLDBERG, A. PASSOW & J. JUSTMAN, *THE EFFECTS OF ABILITY GROUPING* (1966); Heathers, *Groupings*, in *ENCYCLOPEDIA OF EDUCATIONAL RESEARCH* 559 (R. Ebel ed. 1969); Kirp, *Schools as Sorters*, *supra* note 8, at 725-31; Shea, *supra* note 8; Sorgen, *supra* note 8, at 1132-37; *Judicial Supervision*, *supra* note 8, at 1518-22 (and references collected therein); *Educational Testing*, *supra* note 8, at 376-78. There are, however, some important subsequent cases concerning ability grouping. See text accompanying notes 151 to 164 *infra*.

112. *Smuck v. Hobson*, 408 F.2d 175, 189 (D.C. Cir. 1969).

it as referring only to the track system "as it existed at the time of the decree."¹¹³

Although *Hobson* was criticized on several counts,¹¹⁴ from the psychometric point of view the court's gravest error was its insistence that grouping can only be based on tests that measure innate ability.¹¹⁵ No psychologist who has written on the subject believes that tests measure hereditary endowment solely.¹¹⁶ However, there was a second reason for the court's condemnation of group tests. The court asserted that the tests were culturally biased because they were standardized on white middle-class children and thus measured psychological and environmental factors unrelated to the true abilities of black children.¹¹⁷

The court was echoing the universal claim among test critics that standardized tests are heavily biased against poor racial and ethnic minorities because they fail to take into account differences in dialect, value orientation, acquired information, or the importance of the situation or social setting in which tests are administered, and

113. *Id.*

114. *E.g.*, Kurland, *Equal Educational Opportunity: The Limits of Constitutional Jurisprudence Undefined*, 35 U. CHI. L. REV. 595 (1968); *Judicial Supervision*, *supra* note 8.

115. *See, e.g.*, *Ability Grouping*, *supra* note 8, at 160; *Judicial Supervision*, *supra* note 8, at 1519-20; *Intelligence Classifications*, *supra* note 8, at 665 n.89, 671 n.116; *Educational Testing*, *supra* note 8, at 380; 20 STAN. L. REV. 1249 (1968).

116. Cleary, et al., *supra* note 7, at 17; *e.g.*, *id.* at 22: "[I]ntelligence as measured is not a capacity. It is a behavioral trait and one highly dependent on past learning, whether the test is a standard test of intelligence or a 'culture-fair' test." The report from which this quotation was taken, prepared at the request of the American Psychological Association's (APA) Board of Scientific Affairs, was highly criticized by persons representing two minority groups within the APA, but none of the critics argued that the quoted assertion was untrue. Most minority group psychologists believe that current standardized tests measure little, if any, hereditary endowment. Bernal, *supra* note 7; Jackson, *supra* note 7. *See* A. ANASTASI, *supra* note 7, at 350 ("Most intelligence tests can . . . be regarded as measures of scholastic aptitude."); D. HEBB, A TEXTBOOK OF PSYCHOLOGY 246-53 (1958); APA Task Force on Employment Testing of Minority Groups, *Job Testing and the Disadvantaged*, 24 AM. PSYCH. 637, 640 (1969) ("Aptitude tests . . . measure developed abilities, reflecting the interaction between an individual's experiences and his innate endowments."); Anastasi, *Some Implications of Cultural Factors for Test Construction*, in TESTING PROBLEMS IN PERSPECTIVE, 453, 456 (A. Anastasi ed. 1966). Not even Arthur Jensen, the most prolific and publicized of the hereditarians, believes that an intelligence test score reflects innate ability only. A. JENSEN, BIAS IN MENTAL TESTING (1980); Jensen, *How Much Can We Boost IQ and Scholastic Achievement?*, 39 HARV. EDUC. REV. 1 (1969). The argument among psychologists and other scientists concerns the relative weight to be assigned to hereditary and environmental factors. *See generally* R. CANCRO, INTELLIGENCE: GENETIC AND ENVIRONMENTAL INFLUENCES (1971).

117. 269 F. Supp. at 514. Such tests have been called "Anglocentric." J. MERCER, LABELING THE MENTALLY RETARDED 13-15 (1973); Mercer, *supra* note 7, at 138-39.

thus they consistently underpredict the potential ability of those children.¹¹⁸ But, if the assumption in *Hobson* that tests can and should measure innate ability is false, then the condemnation of tests as not reflecting the dominant culture may be misconceived.

The consensus — although not the unanimous view — among psychologists, as reflected in the report of the American Psychological Association Ad Hoc Committee on the Educational Uses of Tests with Disadvantaged Students, is:

[W]hat the tests have measured with greater or lesser validity and comprehensiveness . . . is the possession of abilities that are demonstrated to be useful in predicting future learning. . . . By and large, when properly applied and interpreted, they have predicted future learning for all segments of our society with modest but significant validity and generalizability.¹¹⁹

Thus, tests may be useful in measuring how well students have acquired the skills taught by the school system and how well they may do at the next level in the educational hierarchy.¹²⁰ The tests, rather than reflecting cultural inferiority, are indicative of the educational — not genetic — deficiencies of minority children and,

118. See, e.g., Baca & Cervantes, *The Assessment of Minority Students: Are Adaptive Behavior Scales the Answer?*, 15 PSYCH. IN THE SCHOOLS 366 (1978); Baratz & Baratz, *Negro Ghetto Children and Urban Education: A Cultural Solution*, 33 SOC. EDUC. 401 (1969); Bernal, *supra* note 7; Dent, *Assessing Black Children for Mainstream Placement*, in MAINSTREAMING AND THE MINORITY CHILD 77 (R. Jones ed. 1976); Jackson, *supra* note 7; Jorgensen, *supra* note 7; Labov, *Academic Ignorance and Black Intelligence*, 229 ATLANTIC MONTHLY, June 1972, at 59; Oakland & Matuszek, *Using Tests in Nondiscriminatory Assessment*, in PSYCHOLOGICAL AND EDUCATIONAL ASSESSMENT OF MINORITY CHILDREN 52, 58-62 (T. Oakland ed. 1977); Rivers, Henderson, Jones, Ladner & Williams, *Mosaic of Labels for Black Children*, in 2 ISSUES IN THE CLASSIFICATION OF CHILDREN 213, 231-42 (N. Hobbs ed. 1975); Samuda, *supra* note 7, at 65; *Assessing Minority Group Children* (B. Phillips ed.), 11 J. OF SCH. PSYCH. 291-401 (1973).

119. Cleary, et al., *supra* note 7, at 24. The use of the term "ability" in this widely quoted report is unfortunate because it may perpetuate Judge Wright's confusion concerning what the ideal test should measure, innate ability or present functioning. It would have been more helpful had the authors substituted a term such as "skills," "knowledge," or "overt behaviors" for the word "ability." However, understood in its entirety, the quoted statement can still be accepted as the consensus in the sense that most psychologists agree that tests more or less accurately measure present levels of performance and future learning capabilities. However, insofar as the one standard deviation difference between black and white children on standardized intelligence tests are interpreted by the authors of the Report as racial differences that cannot be ignored, *id.* at 16, the Report has been called "blatantly racist." Jackson, *supra* note 7, at 88.

120. A. ANASTASI, *supra* note 7, at 342.

more important, perhaps the inadequacy of public schools in their present state. "The flaw is in our educational system, from primary school right through college. A tremendous amount of talent is being wasted. . . . [T]ests have been helpful in documenting the severity of the problem."¹²¹ In this respect the "tests are not bigoted villains but color-blind measuring instruments that have demonstrated a social problem to be solved."¹²²

The critics who have been concerned about the impact of low test scores on minority groups and who have repeatedly called for the abolition of testing have themselves been admonished for failing to see the inconsistency of their views and for their failure to argue vigorously enough for changes in the way schools teach:

Low achievement scores reported for groups of minority students have led to demands not for the improvement of the educational system but for the abandonment of those "lying" tests. . . . [W]hen the many carefully constructed and widely used tests are actually inspected for content, almost always that content is found to consist of legitimate samplings of quite uncontroversial educational goals, such as the ability . . . to calculate simple mathematical problems. But the emotionality stemming from the confusion over the proper interpretation of the test scores sometimes leads minority spokespersons to take the position that those tests are asking the *wrong questions*. This implies that somewhere there are some *right* questions and that in fact there is satisfactory achievement taking place on the part of our minority students. . . . It is as if [those spokespersons] are saying that our inner-city schools are in fact being quite successful after all.

But the interesting thing is that some of the same persons who will at least implicitly stand by this argument at some times will at other times be vocal critics of that same school system, claiming that it has failed the minority population. The strain on the logic of the argument occurs when those same low test scores are used to prove the point.¹²³

121. Green, *In Defense of Measurement*, 33 AM. PSYCH. 664, 669 (1978). See Kirp, *Schools as Sorters*, *supra* note 8, at 758.

122. Green, *supra* note 121, at 669. "[E]ven though aptitude tests are culturally biased, they are biased for a reason: the school experience itself is culturally biased." *Ability Grouping*, *supra* note 8, at 161.

123. Flaughner, *supra* note 9, at 672. But see *Martin Luther King Jr. Elem. School Children v. Michigan Bd. of Educ.*, 451 F. Supp. 1324 (E.D. Mich. 1978) (Equal Educational Opportunities Act may guarantee instruction in "Black English" to black students).

The problem lies in defining the nature of the test. If the test is perceived as a measure of achievement, then a low score calls for additional effort on the part of the school to remedy the problem. The bias occurs when the low score is perceived as a measure of pure aptitude. In that case, the low score "may be interpreted as an indication that there is insufficient capacity on that test taker's part to achieve; therefore, any additional educational effort would be wasted."¹²⁴ Insofar as the District of Columbia was failing to provide the educational resources to help low-scoring children compensate for academic deficiencies, the proper remedy should have been improvement of the manner in which the children were taught, not the abolition of the tracking system.¹²⁵

By definition, achievement and intelligence tests will always fail to meet *Hobson's* demand for assessment devices devoid of environmental bias. Given what they purport to measure, they inevitably reflect the social setting of the test taker.

[A]ll behavior is thus affected by the cultural milieu in which the individual is reared and since psychological tests are but samples of behavior, cultural influences will and should be reflected in test performance. It is therefore futile to try to devise a test that is *free* from cultural influences.¹²⁶

Efforts to produce culture-free tests or to reduce content bias have met with little success. "Nonverbal or performance tests are now generally recognized as falling short of the goal of freedom from cultural influences, and attempts to develop culture fair verbal tests . . . are recognized as failures."¹²⁷ "On the WISC [Wechsler Intelligence Scale for Children], for instance, black children usually find the Performance tests as difficult or more difficult than the Verbal tests; this pattern is also characteristic of children from low socioeconomic levels."¹²⁸ "[I]t is sobering but instructive to recognize

124. *Id.*

125. Of course, any recommendation for compensatory education may have to overcome the argument that in-school achievement has little to do with the resources of the school but is rather a reflection of the political, social, and economic environment in which the child lives. See C. JENCKS, *INEQUALITY* (1972).

126. A. ANASTASI, *supra* note 7, at 345 (emphasis in original).

127. Reschly, *supra* note 7, at 231. "The apparent inability of culture-fair tests to yield similar means and standard deviations for persons from different racial-ethnic groups and social classes or to yield suitably high levels of concurrent and predictive validity contributed to their disuse." Oakland & Matuszek, *supra* note 118, at 62.

128. A. ANASTASI, *supra* note 7, at 348 (references omitted).

that minority children do poorly even on so-called culture-free tests"¹²⁹

There has been relatively less research concerning content bias in tests. Criticism focuses on the questions themselves: "They ask questions about things like the distance from Boston to London, how to keep from being lost in the wilderness, the author of Macbeth, the cause of icebergs melting, and such. They ask for the definition of words which are much more frequently used in white than in black culture."¹³⁰ But, contrary to popular thought, such widely criticized questions on the revised version of the WISC (WISC-R) comprehension subtest as, "What is the thing to do if a boy (or girl) much smaller than yourself starts to fight with you?" are actually easier for black children than they are for white.¹³¹ Eliminating thirteen items perceived to be biased from a widely used eighty-two item elementary reading test "did not improve the performance of schools with high minority populations relative to their performance on the original 'biased' version."¹³² Deleting what appear to be idiosyncratic items from group ability tests resulted only in "making the test considerably more difficult for everyone, since many of the items that [show] the widest discrepancy between groups [are] moderate to low in overall difficulty."¹³³

Even if tests are of some use in indicating the erosion of academic skills, they may still be biased with respect to the complementary processes of prediction and selection, the other uses for which supporters believe tests have validity.¹³⁴ Prediction is more

129. Kirp, *Schools as Sorters*, *supra* note 8, at 758. See references collected *id.* at 758 n.265.

130. Jorgensen, *supra* note 7, at 35.

131. Reschly, *supra* note 7, at 232. The implication is that bullying is accepted behavior among blacks. See Kirp, *Schools as Sorters*, *supra* note 8, at 758 n.265.

132. Flaughner, *supra* note 9, at 675.

133. *Id.* Cf. Green, *supra* note 121, at 668 ("Careful attempts to find [culturally biased] items have had slim success. With few exceptions, the same group differences occur for each item."). But see Oakland & Matuszek, *supra* note 118, at 59.

134. See, e.g., BUREAU OF EDUC. FOR THE HANDICAPPED, DEVELOPING CRITERIA FOR THE EVALUATION OF PROTECTION IN EVALUATION PROCEDURES PROVISIONS (1978); Cleary, et al., *supra* note 7; Cleary, *Test Bias: Prediction of Grades of Negro and White Students in Integrated Colleges*, 5 J. EDUC. MEAS. 115 (1968); Cole, *Bias in Selection*, 10 J. EDUC. MEAS. 237 (1973); Cronbach, *Equity in Selection — Where Psychometrics and Political Philosophy Meet*, 13 J. EDUC. MEAS. 31 (1976); Darlington, *Another Look at "Cultural Fairness"*, 8 J. EDUC. MEAS. 71 (1971); Darlington, *A Defense of "Rational" Personnel Selection and Two New Methods*, 13 J. EDUC. MEAS. 43 (1976); Flaughner, *supra* note 9; Hunter & Schmidt, *Critical Analysis of the Statistical and Ethical Implications of Various Definitions of Test Bias*, 83 PSYCH. BULL. 1053 (1976); Kennedy, *Rationality, Emotionality, and Testing*, 16 J. SCH. PSYCH. 16, 19-22 (1978); Linn & Werts, *Considerations for Studies of Test Bias*, 8

pertinent to the topic of selection for higher education and employment, and extensive discussion would be beyond the scope of this Article. Several points, however, should be made.

First, the issue of cultural bias in selection based on test scores is complex and controversial, and the opinions on it contradictory,¹³⁵ although there is some consensus that tests cannot be fair unless they predict with equal accuracy for all groups.¹³⁶ Second, studies investigating the ability to predict academic achievement from standardized tests indicate that "minority groups and people of

J. EDUC. MEAS. 1 (1971); McNemar, *On So-Called Test Bias*, 30 AM. PSYCH. 848 (1975); Novick & Ellis, *Equal Opportunity in Educational and Employment Selection*, 32 AM. PSYCH. 306 (1977); Novick & Petersen, *Towards Equalizing Educational and Employment Opportunity*, 13 J. EDUC. MEAS. 77 (1976); Oakland, *Predictive Validity of Readiness Tests for Middle and Lower Socioeconomic Status of Anglo, Black and Mexican American Children*, 70 J. EDUC. PSYCH. 574 (1978); Petersen & Novick, *An Evaluation of Some Models for Culture-Fair Selection*, 13 J. EDUC. MEAS. 3 (1976); Sawyer, Cole & Cole, *Utilities and the Issue of Fairness in a Decision Theoretic Model for Selection*, 13 J. EDUC. MEAS. 59 (1976); Schmidt & Hunter, *Racial and Ethnic Bias in Psychological Tests: Divergent Implications of Two Definitions of Test Bias*, 29 AM. PSYCH. 1 (1974); Thorndike, *Concepts of Culture-Fairness*, 8 J. EDUC. MEAS. 63 (1971); Reschly, *Comparison of Bias in Assessment with Conventional and Pluralistic Measures* (unpublished paper presented at the Annual Convention of the Council for Exceptional Children) (May 1978).

135. See, e.g., Petersen & Novick, *supra* note 134, at 23-24:

[The many models for test bias] are each explications of general concepts of what constitutes fair use of tests in a selection situation. There seems to be nothing in the literature that clearly indicates when, if ever, one of the models is preferable to the other five models. Thus, the practitioner has no clear guidance in the choice of a culture-fair selection model. Further, we have suggested that [several models] and their converses are *internally contradictory*.

As stated in Ysseldyke, *Implementing the "Protection in Evaluation Procedures" Provisions of Public Law 94-142*, in BUREAU OF EDUC. FOR THE HANDICAPPED, DEVELOPING CRITERIA FOR THE EVALUATION OF PROTECTION IN EVALUATION PROCEDURES PROVISIONS 150 (1978):

Several investigators have reviewed the models of test fairness and have concluded that there is little agreement among the several models. It is readily apparent that major measurement experts have been essentially *unable* to agree on a definition of a fair test, let alone identify a test that is fair for members of different groups. There is little agreement on the *concept* of nondiscriminatory assessment.

See Flaughner, *supra* note 9, at 671: "[A]spects [of test bias] are widely disparate and frequently stem from entirely different universes of discourse"; Schmidt & Hunter, *supra* note 134, at 1; cf. Hunter & Schmidt, *supra* note 134, at 1069 (ethical positions regarding test bias were shown "to be irreconcilable").

136. Reschly, *supra* note 7, at 233-34. This position must be distinguished, however, from the one that asserts that tests are only fair if there are no mean score differences between population subgroups. That position is considered "untenable" in that "such a definition eliminates a priori any possibility of real group differences on various psychological traits . . ." Schmidt & Hunter, *supra* note 134, at 1.

lower SES [socioeconomic status] are generally predicted just as well as are upper-middle SES Caucasians."¹³⁷ The WISC-R Full Scale I.Q. Score predicts school achievement similarly for white, black, and Chicano elementary and junior high school children.¹³⁸ Third, when test bias is measured by certain forms of statistical analysis, the result is that in some situations the standardized test overpredicts success for minorities and underpredicts success for whites.¹³⁹ Fourth, other definitions of test bias, which seek to correct disproportionate selection procedures by statistical adjustments in test scores, cutoff points, and prediction formulas or by establishing quotas, have not only met with disapproval by those in measurement¹⁴⁰ but may run afoul of *Bakke's* proscription of the use of explicit and purposeful racial criteria for selection purposes.¹⁴¹ Finally, and perhaps most important, reliance on psychometric models of test bias without consideration of the social and ethical consequences of test use ignores the concerns of significant segments of society. While the Ad Hoc Report¹⁴² defended the technical adequacy of tests for prediction and selection, it failed to

137. Kennedy, *supra* note 134, at 19-20.

138. Reschly & Reschly, *Validity of WISC-R Factor Scores in Predicting Achievement and Attention for Four Sociocultural Groups*, 16 J. SCH. PSYCH. (in press). However, the WISC-R predicted less well for native American children. The correlations themselves fell within the usual range of .40 to .60 found between I.Q. and school achievement. Correlation between the WISC-R and Reading Scores on the Metropolitan Achievement Test were .56, .62, and .55 for white, black, and Chicano children respectively. Correlations between the I.Q. test and teacher ratings as measured by a ten-item Academic Scale were .35, .45, and .38 for the same groups. In *Larry P. v. Riles*, No. C-71-2270 R.F.P. (N.D. Cal. Oct. 16, 1979), *appeal docketed*, No. 80-4027 (9th Cir. Jan. 17, 1980) (*Riles II*), the court rejected such studies to support the validity of I.Q. tests for minorities. The court would only consider studies in which I.Q. test results were correlated with classroom grades, not with other standardized tests. Slip op. at 71. *Riles II* is discussed in the text accompanying notes 279 to 351 *infra*.

139. The principal method used in this computation is called "regression analysis." It permits researchers to predict one characteristic of an individual from one or more other characteristics. For example, school achievement may be predicted from intelligence test scores. J. WERT, C. NEIDT & J. AHMANN, *STATISTICAL METHODS IN EDUCATIONAL AND PSYCHOLOGICAL RESEARCH* 226 (1954). The most widely followed discussion of the subject is in Cleary, *Test Bias: Prediction of Grades of Negro and White Students in Integrated Colleges*, 5 J. EDUC. MEAS. 115 (1968). The technique there described was adopted by the APA Ad Hoc Committee on Educational Uses of Tests with Disadvantaged Students. Cleary, et al., *supra* note 7, at 25-32. See A. ANASTASI, *supra* note 7, at 192.

140. E.g., A. ANASTASI, *supra* note 7, at 197; Hunter & Schmidt, *supra* note 134. For criticism of those techniques (which the authors call "group parity models") by a psychometrician and legal scholar, see Novick & Ellis, *supra* note 134.

141. Regents of Univ. of Cal. v. Bakke, 438 U.S. 265 (1978).

142. Cleary, et al., *supra* note 7.

consider what minority groups charged was the egregious misuse of tests having a negative impact on the lives of minorities.¹⁴³ "[T]o defend tests on the basis of evidence of common regression systems or to attempt to separate the issues of technical adequacy from those of social consequences is insufficient."¹⁴⁴ Recent attempts to examine the ethical, legal, and social implications of various models of test bias are valuable additions to the literature.¹⁴⁵ In essence, even the selection of a model to measure and ameliorate test bias is ultimately a value judgment.¹⁴⁶ Justice Powell, writing in *Bakke*, appears to have opted for what may be called "qualified individualism,"¹⁴⁷ i.e., that considerations of race and ethnicity may be appropriate under some circumstances but not when selection is based solely on those characteristics.¹⁴⁸

Aftermath of Hobson

A significant finding by Judge Wright in *Hobson* was that reliance on group ability tests contributed to the misclassification of approximately 820 of 1272 students.¹⁴⁹ Evidence of the misclassification was provided by the school system itself. In 1965, two years prior to the *Hobson* decision, the school superintendent had ordered that no student could be assigned to the EMR track without an evaluation, usually an individual test by a psychologist. When clinicians reassessed the children in the Special Academic Track they found that almost two-thirds were not genuinely retarded. In this light, *Hobson* could be seen as a vindication of the use of individual tests by school and clinical psychologists.¹⁵⁰ Nevertheless,

143. Bernal, *supra* note 7; Jackson, *supra* note 7.

144. Reschly, *supra* note 7, at 235.

145. See, e.g., Hunter & Schmidt, *supra* note 134; Novick & Ellis, *supra* note 134.

146. For example, one may choose a group-parity model so as to reduce the effects of past discrimination or mitigate current skill deficits in minority groups. Or, one may choose a model that insures that the primary consideration is equal opportunity for individuals, regardless of race. Both may be considered to have high social utility, but each will have different consequences for minorities.

147. The term was coined by Hunter & Schmidt, *supra* note 134, at 1054, but used for different purposes.

148. "Fairness in individual competition for opportunities, especially those provided by the State, is a widely cherished American ethic." Regents of Univ. of Cal. v. Bakke, 438 U.S. 265, 319 n.53 (1978); see Hunter & Schmidt, *supra* note 134, at 1053-54, 1066-69.

149. *Hobson v. Hansen*, 269 F. Supp. 401, 490 (D.D.C. 1967), *aff'd sub nom.* Smuck v. *Hobson*, 408 F.2d 175 (D.C. Cir. 1969). See also 408 F.2d at 197.

150. Of course, there were no data that validated the determinations of the clinicians based on those individual tests. The tests themselves were not identified in the opinion or in the affirmance.

after *Hobson* both ability grouping and testing of all kinds came under intense judicial scrutiny.

In 1970 a second series of cases challenging the pupil placement process began with *Singleton v. Jackson Municipal Separate School District*.¹⁵¹ Title VI of the Civil Rights Act¹⁵² had been in effect for six years, and, faced with southern intransigence,¹⁵³ the Supreme Court declared that freedom of choice plans that failed to eliminate racial discrimination "root and branch" were unconstitutional¹⁵⁴ and that school boards were "obligated to terminate dual school systems at once and to operate [unitary schools] now."¹⁵⁵ In *Singleton*, without explanation and without exploration of the validity of either ability grouping or testing, the Fifth Circuit held that achievement testing for the purpose of assigning students to schools could not "be employed . . . until unitary school systems have been established."¹⁵⁶ This holding was expanded in *Lemon v. Bossier Parish School Board*,¹⁵⁷ a case in which the court found that a semester of desegregation was not long enough to permit a school system to consider standardized testing for the purpose of pupil assignment. The Fifth Circuit again declined to rule on the validity of testing per se even though it had been asked to do so by the black plaintiffs. It simply stated that "when a school district . . . has operated as a unitary system for a sufficient time . . . we will then decide that complex and troubling question" of test validity.¹⁵⁸

The issue was more refined in *Moses v. Washington Parish School Board*,¹⁵⁹ in which a federal district court explored the constitutionality of ability grouping *within* schools. The school

151. 419 F.2d 1211 (5th Cir. 1970), *rev'd in part on other grounds sub nom. Carter v. West Feliciana School Bd.*, 396 U.S. 290 (1970), *vacated in part*, 396 U.S. 226 (1969). The case decided by the court of appeals was a consolidation of 17 separate district court cases. Certiorari to the Supreme Court was denied in two of those cases. 396 U.S. 1032 (1970).

152. 42 U.S.C. § 2000d (1974): "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving federal financial assistance."

153. From 1954 to 1964 only 2.14% of black students attended desegregated schools in the 11 southern states. D. KIRP & M. YUDOF, *supra* note 8, at 307.

154. *Green v. County School Bd.*, 391 U.S. 430, 438 (1967).

155. *Alexander v. Holmes County Bd. of Educ.*, 396 U.S. 19, 20 (1969).

156. 419 F.2d at 1219. Two identical rulings followed in *United States v. Sunflower County School Dist.*, 430 F.2d 839 (5th Cir. 1970), and *United States v. Tunica County School Dist.*, 421 F.2d 1236 (5th Cir. 1970).

157. 444 F.2d 1400 (5th Cir. 1971).

158. *Id.* at 1401.

159. 330 F. Supp. 1340 (E.D. La. 1971), *aff'd*, 456 F.2d 1285 (5th Cir. 1972).

board, upon being ordered to integrate, decided to create homogeneous classes in all schools based solely on reading achievement. The expert witnesses who testified at trial all agreed that if homogeneous grouping was to function soundly, classification should be based on multidimensional assessment and that assignment for all subjects grounded only on reading test scores was "the worst form of grouping."¹⁶⁰ As in *Hobson*, the court found that the program was not genuinely compensatory,¹⁶¹ that there was little between-track fluidity, and that the lower tracks were almost totally populated by black students. Without determining the per se validity of standardized tests, and after acknowledging that "courts are not school boards and do not derogate unto themselves the formulation of educational policies,"¹⁶² the court held that assignment of black students "on the basis of presently used testing" violated the equal protection clause and enjoined the use of such testing.¹⁶³ For the first time in this series of cases, the court provided a rationale for its ruling. Though it condemned neither ability grouping nor testing, it found the assessment process discriminatory against black students who had been "educated in admittedly inferior [and segregated] schools [and who were now called upon to compete with] white students educated in superior schools for position in top sections."¹⁶⁴

160. *Id.* at 1343.

161. The school system claimed that although it did not provide any added instruction, the homogeneous grouping plan was compensatory by permitting slow students to be placed with other slow students and to be exposed to regular school materials at a slower pace. *Id.*

162. *Id.* at 1345.

163. *Id.*; "Testing used to resegregate students in a recently desegregated school system is not permissible." *Id.*

164. *Id. Accord, McNeal v. Tate County School Dist.*, 508 F.2d 1017, 1019 (5th Cir. 1975). This view has been criticized:

The evidence is conflicting as to whether blacks were psychologically or academically harmed by segregation. It is also difficult to argue that black children attending school in 1972 are victims of the discrimination of the 1950's directed against other black children. Even assuming a lingering injury, there is conflicting evidence as to whether integration will undo that injury. Nonetheless, in the absence of persuasive evidence to the contrary, courts have steadfastly rejected ability grouping as an adequate response to the *Brown* mandate.

Yudof, *supra* note 8, at 451. *But see Kirp, Schools as Sorters, supra* note 8, at 763. This criticism, of course, is only relevant where black plaintiffs challenging tests and tracking were not themselves the victims of segregated education in second-rate schools.

Ironically, in *McNeal v. Tate County School Dist.*, 508 F.2d 1017, 1019 (5th Cir. 1975), a case in which the Fifth Circuit approved the analysis in *Moses*, ability grouping was determined not by standardized tests, but by teacher evaluations of pupil performance. Tracking was barred unless the school district could validate its system. Thus, it did not seem to matter what the particular scheme was in these cases;

The Ninth Circuit, unlike the Fifth, looked to *Hobson* for authority but, in contrast to *Hobson*, found *individual testing* unsatisfactory.¹⁶⁵ The first Ninth Circuit case was *Diana v. State Board of Education*,¹⁶⁶ a class action in which nine representative Mexican-American elementary school students challenged their placement in EMR classes. Their school district classified pupils as

what concerned the courts was the use of apparently racially neutral methods of classification in a discriminatory fashion. Tests were subsumed in the general condemnation of the school districts' attempts to maintain segregation through subterfuge.

165. There were, however, two district court cases arising in the Ninth Circuit that resembled their Fifth Circuit counterparts. In *Spangler v. Pasadena City Bd. of Educ.*, 311 F. Supp. 501 (C.D. Cal.), *aff'd*, 427 F.2d 1352 (9th Cir. 1970), *cert. denied*, 402 U.S. 943 (1971), the district court found that interclass grouping within integrated schools based in part on achievement and intelligence test scores violated the fourteenth amendment. The defendant school system agreed that those tests were "racially discriminatory, based as they are primarily on verbal achievement." *Id.* at 519. However, "because of the delicate educational nature of decisions concerning grouping" the court only urged the defendant to reexamine its grouping policies carefully. *Id.* at 504. In *Johnson v. San Francisco Unified School Dist.*, 339 F. Supp. 1315 (N.D. Cal. 1971), *vacated and remanded*, 500 F.2d 349 (9th Cir. 1974), the court enjoined San Francisco from "authorizing, permitting or using tracking systems or other educational techniques or innovations without effective provisions to avoid segregation." *Id.* at 1325. In *Berkelman v. San Francisco Unified School Dist.*, 501 F.2d 1264 (9th Cir. 1974), however, a Ninth Circuit panel held that it was constitutional for a college preparatory high school to admit only those students whose academic achievement placed them in the top 15% of their junior high school class even though this excluded a disproportionate number of black and Spanish-American students. Distinguishing *Hobson*, the court asserted that the unadmitted students who thus had to attend the regular comprehensive high schools were not denied a quality education. Further, the court noted that the classification system in question depended upon past achievement, impartially measured, rather than on test-based predictions that isolated students of lower ability. Finally, the court found that the plaintiff students would suffer little psychic injury in attending schools in which the majority of city students were enrolled. While it was bothered by the racial and ethnic disproportionality, "uncertainty on this score . . . does not mean that the maintenance of [the elite school] is itself unconstitutional or that conditioning admission on the basis of past academic achievement is unconstitutional." *Id.* at 1268. The reasoning of the court with regard to selection based on past versus predicted achievement is somewhat specious. In both cases, the evaluation data were being used to predict future behavior. While it is true that in the one case the data forecast high ability, and in the other, lack of ability to do even grade appropriate work, here, like *Hobson*, past achievement was used to predict future achievement. Implied also in *Berkelman* is the judgment that grade point averages are more objective and valid than are standardized achievement tests. While grades themselves may be based on subjective and irrelevant evaluations, even test advocates acknowledge that "the best single predictor of a future academic record is usually the past academic record. . . ." Cleary, et al., *supra* note 7, at 34; *accord*, L. CRONBACH, *supra* note 7, at 289.

166. C.A. No. C-70-37 R.F.P. (N.D. Cal., filed Feb. 3, 1970). The case led to a stipulated agreement and thus no decision was reported. For copies of the complaint and the consent decree, see CLASSIFICATION MATERIALS, *supra* note 8, at 199-229.

educably retarded if they scored between 55 and 70 on the WISC or between 52 and 68 on the Stanford-Binet.¹⁶⁷ The plaintiffs contended that the tests were unrelated to their ability to learn, culturally biased, too dependent on verbal skills, improperly standardized (only whites constituted the norm group), and inappropriately administered because Spanish-speaking children were tested solely in English by a non-Spanish-speaking examiner. The outcome was significant ethnic disproportionality in the county's EMR program; Spanish surnamed students constituted a little over eighteen percent of the student population but accounted for nearly one-third of the children in EMR classes.¹⁶⁸

The plaintiffs' most damaging evidence was provided by a bilingual psychologist who retested the children. When reexamined in Spanish and English, Diana, the named plaintiff, achieved an I.Q. score of 79, forty-nine points higher than when initially tested. Seven of the other eight plaintiffs also scored at or above the cutoff score the district had set for EMR placement.¹⁶⁹ In the face of these data, the California Department of Education agreed to a consent decree that stipulated the following changes in its method of evaluating bilingual children:¹⁷⁰ (1) all children whose primary language was not English would be tested in both their primary language and in English; (2) unfair verbal items — certain vocabulary and informational questions, for example — would be eliminated; (3) all Mexican-American and Chinese children in EMR classes would be reevaluated in their primary language and only as to those items that measured nonverbal skills; and (4) the state

167. The difference is accounted for by the difference in standard deviations for the tests, 15 points on the WISC and 16 points on the Stanford-Binet. In both instances, however, the criterion would be two to three standard deviations below the mean of 100.

168. Similar results were found in California as a whole. See *CLASSIFICATION MATERIALS*, *supra* note 8, at 203; *Cultural Bias*, *supra* note 8, at 1033 n.21.

169. The remaining child had an I.Q. score of 67 on the WISC. The range of the nine scores was 67 to 89, derived from the WISC. The mean was approximately 77. The California State Department of Education's own study in 1969 produced similar results. Of 47 Mexican-American children in urban and rural school districts, 42 who were retested on the Spanish version of the WISC scored above the state-mandated cutoff point. The mean difference was 13.15 points (68.61 to 81.76). Twenty-seven of the children scored above 80. California State Dep't of Education, *Spanish-Speaking Pupils Classified as Educable Mentally Retarded*, 7 *INTEGRATED EDUC.* 29-31 (1969).

170. *CLASSIFICATION MATERIALS*, *supra* note 8, at 226. The state also agreed that counties would provide supplemental education so that EMR-placed children could be reintegrated into regular classes. In addition, any school district that had a significant disparity of Mexican-American children in its regular and EMR classes would have to justify that disparity. *Id.*

would develop revised I.Q. tests reflecting Mexican-American culture and normed only on Mexican-American children.

Although it is difficult, under any circumstances, to accomplish the second item¹⁷¹ and although the fourth has never been accomplished, the rationale behind the decree is sound. Testing non-English-speaking or bilingual children in a language they do not fully understand and, on that basis, labeling them as retarded, is recognized as an egregious example of cultural bias.¹⁷² Perhaps the most far-reaching outcome of *Diana* was that major provisions of the consent decree were incorporated into the California Education Code.¹⁷³

In *Covarrubias v. San Diego Unified School District*,¹⁷⁴ the testing issue was broadened to include charges of racial, cultural, and environmental bias as well as inappropriate administration of I.Q. testing for EMR placement. Like *Diana*, *Covarrubias* was settled by a stipulated agreement, but in addition to seeking to enjoin the continued use of I.Q. tests the black and Mexican-American plaintiffs asked for \$400,000 in punitive damages for the district's

171. See text accompanying notes 127 to 133 *supra*.

172. See, e.g., Cleary, et al., *supra* note 7, at 21-22, who call this "test misuse," and Flaughner, *supra* note 9, at 677, who sees it as an example of "atmosphere bias." Use of an intelligence test in this manner has been distinguished from a use in which the purpose is to assess a child's current status in understanding English and ability to manipulate stimuli thought important by the dominant culture. See Cleary, et al., *supra* note 7, at 22.

173. By 1971 the legislature had enacted laws declaring the following: (1) there should be no disproportionality in EMR classes with regard to socioeconomic, racial, or ethnic status, CAL. EDUC. CODE §56504 (West 1978); (2) all intelligence tests administered for the purpose of placement in EMR programs should be given in the language in which the child is most fluent, *id.* § 56505; (3) no child should be placed in an EMR class if he or she scores higher than two standard deviations below the norm on a nonverbal intelligence test (or nonverbal portion of a verbal/nonverbal test) unless an admissions committee unanimously agrees in writing after considering all pertinent information that the child belongs in a class for retarded students, *id.* § 56506; (4) no child should be placed in an EMR class unless given a comprehensive evaluation by a properly certified school psychologist, including not only an intelligence test but data from a developmental history, educational evaluation, and adaptive behavior scale, *id.* § 56508; (5) no evaluation or placement should take place without written consent of the parents, *id.* § 78804-05. With regard to § 56508, adaptive behavior is a measure of the extent to which the person can play "a full complement of social roles appropriate for his age and is performing in those roles in a manner comparable to that of other persons of his age in the society." J. MERCER, *supra* note 117, at 137. In 1974 the state was ordered by the court to undo all disproportionality in EMR classes with regard to Mexican-American children. Kirp, Kuriloff & Buss, *Legal Mandates and Organizational Change*, in 2 ISSUES ON THE CLASSIFICATION OF CHILDREN 319, 364 (N. Hobbs ed. 1975).

174. Civ. No. 70-394-S (S.D. Cal., filed Feb. 1971) (settled by consent decree, July 31, 1972).

mislabeling and misplacement of children in EMR classes. Although the plaintiffs were awarded only nominal damages, the case set a precedent by establishing the appropriateness of both equitable relief and monetary damages in misclassification litigation.¹⁷⁵ A third testing case in this circuit, *Guadalupe Organization, Inc. v. Tempe School District No. 3*,¹⁷⁶ also ended in a settlement in which the school system agreed to reevaluate all children placed in EMR classes, to test bilingual children in their primary language, and to use more information than intelligence test scores in placement decisions.¹⁷⁷

Larry P. v. Riles: Preliminary Injunction Phase

If *Hobson* was the seminal case of the 1960's, *Larry P. v. Riles*¹⁷⁸ deserves similar status for the 1970's. The trial court's decision on the merits, which took eight years to reach, threatens the continued administration of individual intelligence tests and the existence of EMR classes, particularly as they involve minority children. As the rationale of the decision will almost certainly guide future litigation concerning psychological assessment, the case warrants detailed examination. The case has had two phases — the granting of a

175. In a similar misclassification suit, *Stewart v. Phillips*, C.A. No. 70-1199-F (D. Mass., filed Sept. 14, 1970), plaintiffs asked for \$20,000 in compensatory damages for each child improperly placed in EMR classes in the Boston school system. The suit triggered the passage of a new special education act in Massachusetts, MASS. GEN. LAWS ANN. ch. 71B, §§ 1-14 (West Supp. 1978). As of February 1980 the case was finally being prepared for trial.

176. Civ. No. 71-435 (D. Ariz., filed August 9, 1971) (settled by consent decree, Jan. 24, 1972). In a subsequent action, plaintiffs' claim that the school's bilingual instructional program violated the equal protection clause of the fourteenth amendment, Title VI of the 1964 Civil Rights Act, 42 U.S.C. § 2000d (1970), and the Equal Education Opportunity Act of 1974, 20 U.S.C. § 1703 (1974), was rejected. *Guadalupe Org'n, Inc. v. Tempe Elem. School Dist. No. 3*, 587 F.2d 1022 (9th Cir. 1978).

177. The most important case concerning bilingual education is *Lau v. Nichols*, 414 U.S. 563 (1974), in which Chinese students charged that the San Francisco school system failed to provide bilingual language instruction to all children in need of such programs. The Supreme Court agreed, relying solely on Title VI of the 1964 Civil Rights Act, 42 U.S.C. § 2000d (1970), and the regulations promulgated under that provision, 45 C.F.R. § 80 (1974). By basing its decision on nonconstitutional grounds, the Court not only reinforced the right of bilingual children to a "meaningful education," 414 U.S. at 566, but perhaps more significantly, gave its imprimatur to both Title VI and HEW's regulations, thus spurring the development of other regulations bearing on the issue of psychological testing.

178. 343 F. Supp. 1306 (N.D. Cal. 1972) (order granting preliminary injunction), *aff'd*, 502 F.2d 963 (9th Cir. 1974), No. C-71-2270 R.F.P. (N.D. Cal. Oct. 1979) (decision on the merits), *appeal docketed*, No. 80-4027 (9th Cir. Jan. 17, 1980).

preliminary injunction in 1972 (*Riles I*),¹⁷⁹ and the decision on the merits in 1979 (*Riles II*).¹⁸⁰

In 1971 black children attending San Francisco schools filed suit in federal district court charging discrimination in their placement in EMR classes as a result of having scored lower than 75 on state-approved intelligence tests. The plaintiffs claimed they were not mentally retarded and that the tests used to place them were culturally biased. They alleged that the resultant classification violated the equal protection clause and requested a preliminary injunction restraining the school system from using I.Q. tests to determine EMR placement of black children.

Adopting the approach used in *Diana v. State Board of Education*,¹⁸¹ the plaintiffs presented affidavits from several black psychologists who had retested the children. Although they administered the identical tests initially given plaintiffs, the psychologists did so only after they made attempts to establish rapport, took pains to reduce distraction, and reworded items in language considered more consistent with the children's cultural background. Scoring procedures were changed so that the children were given credit for nonstandard answers that were judged to show an intelligent approach to solving the problem. The consequence of these efforts was that on retesting all of the plaintiffs scored above the 75 I.Q. cutoff point.

To show the irreparable injury required for the issuance of a preliminary injunction,¹⁸² the plaintiffs made several assertions. Injury was claimed to be caused by plaintiffs' placement in a program in which the curriculum was poor, teacher expectations low, ridicule from peers high, and where placement was permanently noted in the school's cumulative records, creating a stigma that might affect future schooling, employment, and military enlistment.¹⁸³ The school system defended the EMR program by pointing to its relatively easier curriculum, slower pace, lack of pejorative labels — the classes were characterized "ungraded" or "adjustment"

179. 343 F. Supp. 1306 (N.D. Cal. 1972), *aff'd*, 502 F.2d 963 (9th Cir. 1974).

180. No. C-71-2270 R.F.P. (N.D. Cal. Oct. 16, 1979), *appeal docketed*, No. 80-4027 (9th Cir. Jan. 17, 1980). See text accompanying notes 279 to 351 *infra*.

181. C.A. No. C-70-37 R.F.P. (N.D. Cal., filed Feb. 3, 1970). See text accompanying notes 166 to 173 *supra*.

182. See generally O. FISS, INJUNCTIONS 168 (1972).

183. Plaintiffs also charged that harm would be especially great since reevaluation of their intellectual status occurred only once every three years. However, between the time of filing and the consideration of injunctive relief, California's code was revised to require annual reevaluation. CAL. EDUC. CODE § 56501 (West 1978).

— and the opportunity for reintegration into the mainstream on the basis of yearly evaluations.

The court rejected the defendant's position, concluding that "even if a student remains in an EMR class for only one month, that placement is noted on his permanent record, his education is retarded to some degree and he is subjected to whatever humiliation students are exposed to for being separated into classes for the educable mentally retarded."¹⁸⁴ In the court's view, therefore, any wrongful placement would constitute irreparable harm.

The key aspect of *Riles I* was the court's determination of the equal protection issue. The plaintiffs contended that because the testing practices of the San Francisco school system disproportionately harmed black children, their use by the defendants violated the equal protection clause of the fourteenth amendment.¹⁸⁵ They did not

184. 343 F. Supp. at 1308.

185. The equal protection clause provides that no state shall "deny to any person within its jurisdiction the equal protection of the laws." U.S. CONST. amend. XIV, § 1. It comes into play only when the alleged inequality is the result of state action, as are all actions of public schools for fourteenth amendment purposes. See *Brown v. Board of Educ.*, 347 U.S. 483 (1954). It is not sufficient for a plaintiff to show that a particular practice treats different individuals differently, *Griffin v. County School Bd.*, 377 U.S. 218 (1964), or that it fails to achieve equality with mathematical precision, *Lindsley v. Natural Carbonic Gas Co.*, 220 U.S. 61 (1911). Rather, the courts will only hold classifications unconstitutional if they find them to be "invidious" or "arbitrary" and only after examining the circumstances in each case, the nature of individual interests involved, the purported state interest at stake, and the particular groups affected. See, e.g., *Kramer v. Union Free School Dist.*, No. 15, 395 U.S. 621, 626 (1969).

In the main, courts have used two standards or tests when a state statute or practice is called into question as offending the equal protection clause. The usual standard has been designated as the rational basis test or the relaxed standard of review. Under this test a state may differentiate among persons as long as the classification bears some rational relationship to a legitimate state purpose. *McGowan v. Maryland*, 366 U.S. 420, 425 (1961). The state may justify its classificatory scheme by showing that it reasonably seeks to fulfill the purpose for which it is developed. This constraint does not impose a significant burden upon the state. In practice, there is considerable deference to legislative prerogative. Courts will not only presume that the state had a reasonable basis for enacting a particular measure, but will often hypothesize the government's purpose behind its classification mechanism in cases in which it fails to state its rationale explicitly. E.g., *Williamson v. Lee Optical Co.*, 348 U.S. 483, 489 (1955). Moreover, courts have traditionally placed the burden on the plaintiff to show that the state's action is arbitrary and unreasonable. *McGowan v. Maryland*, 366 U.S. 420 (1961). The second test, commonly called the compelling state interest test or the strict scrutiny standard, is much more difficult for the state to meet. When applied, the state must show a compelling or overriding reason for its practice, the practice must be precisely tailored to accomplish its purpose, and there must be no less drastic means for fulfilling its objective. *Dunn v. Blumstein*, 405 U.S. 330 (1972); *Shelton v. Tucker*, 364 U.S. 479 (1960). The strict scrutiny test is triggered in either of two cases — when the state's classification scheme impairs the exercise of a fundamental right or when it affects certain groups

contend that education was a fundamental right of which they were deprived, nor did they assert that the classification of black children as retarded was discrimination based explicitly on race. Rather, they asserted that although placement in EMR classes was accomplished through a purely intelligence-based classification, the method of classification led to a disproportionate impact on black children. It was undisputed that although blacks constituted only 28.5 percent of all students in the San Francisco school system, 66 percent of all students in the EMR program were black. Similarly, though blacks comprised 9.1 percent of the California school population, 27.5 percent of all school children in the state in EMR classes were black. The plaintiffs urged rejection of traditional equal protection analysis, which forced them to shoulder the burden of proving that the classification process was arbitrary and irrational. They claimed

known as "suspect classes." Fundamental rights are those either specifically guaranteed by the Bill of Rights — e.g., speech, press, religious expression, and trial by jury — or those not expressly stated but which are considered critical to the preservation of basic civil and political rights — e.g., interstate travel, privacy, and voting. With regard to the suspect class, it is not meant that the group has suspicious attributes but rather that the classification is suspect because it unduly burdens certain carefully defined categories of people considered to be "saddled with such disabilities, or subjected to such a history of purposeful unequal treatment, or relegated to such a position of political powerlessness as to command extraordinary protection from the majoritarian political process." *San Antonio Ind. School Dist. v. Rodriguez*, 411 U.S. 1, 28 (1973); see *United States v. Carolene Prods. Co.*, 304 U.S. 144, 152 n.4 (1938). To date, only national ancestry, alienage, and race are considered suspect classes. The apparent bases for the denomination of these groups as suspect is that the characteristics they possess are congenital and essentially immutable, that classifications affecting them stigmatize the class in the eyes of society, and, through no fault of their own, they are, or have been, politically impotent. See *Intelligence Classifications*, *supra* note 8, at 653-54. Significant consequences flow from the Supreme Court's description of a right as fundamental or a class as suspect. Either identification has been tantamount to a court's finding that the challenged practice violates the equal protection clause. "[T]he determination of whether someone needs the special protection of the compelling interest standard thus involves complex policy evaluation. The Court, therefore, has been noticeably reluctant to expand the range of fundamental interests and suspect classifications." Note, *Constitutional Requirements for Standardized Ability Tests in Education*, *supra* note 8, at 804. See text accompanying notes 227 to 231 *infra*.

Recently, the Supreme Court appears to have developed an intermediate standard of review, requiring states to show "some ground of difference having a fair and substantial relation to the object of the legislation." *Reed v. Reed*, 404 U.S. 71, 76 (1971). This test has been employed most often when the Court has scrutinized practices that have allegedly discriminated between the sexes, e.g., *Craig v. Boren*, 429 U.S. 190, 197 (1976); see Redish, *Legislative Response to the Medical Malpractice Insurance Crisis: Constitutional Implications*, 55 TEX. L. REV. 759, 773-79 (1977), but it has also been used where a non-suspect classification operates to the detriment of a disadvantaged class of persons, *De La Cruz v. Tormey*, 582 F.2d 45, 59 n.10 (9th Cir. 1978); *Berkelman v. San Francisco Unified School Dist.*, 501 F.2d 1264, 1267 (9th Cir. 1974).

that the burden should shift instead to the defendant school system to demonstrate the rationality of its procedures.

The plaintiffs had strong precedential support in cases of an analogous nature. In the past, courts had shifted the burden when qualification tests for jury service led to disproportionately low numbers of blacks on grand and petit juries¹⁸⁶ and when school board policies with regard to delineation of school boundaries resulted in disproportionate predominance of one race within schools.¹⁸⁷ Perhaps the most persuasive and apt precedent for a shift in burden was *Griggs v. Duke Power Co.*,¹⁸⁸ a case decided by the Supreme Court the year before *Riles I*. There, in an employment discrimination action brought under Title VII of the 1964 Civil Rights Act,¹⁸⁹ black employees challenged the use of intelligence tests as a condition of employment or transfer to certain positions for which they were otherwise qualified. The employer claimed that although the test criterion may have had a discriminatory effect in that fewer blacks were hired or promoted, it had had no intent to discriminate. The Supreme Court, however, interpreted Title VII as proscribing "not only overt discrimination but also practices that are fair in form, but discriminatory in operation."¹⁹⁰ Despite the lack of any evidence that promotion policies were designed to prevent the advancement of black employees — to the contrary, the court found some evidence of special efforts by the employer to aid blacks — the Court declared that "good intent or absence of discriminatory intent does not redeem . . . testing mechanisms that operate as 'built-in headwinds' for minority groups. . . ."¹⁹¹ Once the discriminatory effect was shown, the Court placed the burden in Title VII cases on the employer to show "that any given requirement . . . ha[d] a manifest relationship to the employment in question."¹⁹²

186. *E.g.*, *Carmical v. Craven*, 457 F.2d 582 (9th Cir. 1971).

187. *E.g.*, *United States v. School Dist. 151*, 286 F. Supp. 786 (N.D. Ill.), *aff'd*, 404 F.2d 1125 (7th Cir. 1968).

188. 401 U.S. 424 (1971).

189. 42 U.S.C. § 2000e-2 (1974). Title VII prohibits discrimination in employment because of race, color, sex, or national origin. In 1972 the Act was amended to cover employment by state and local governments. Equal Employment Opportunity Act of 1972, Pub. L. No. 92-261, § 2(2), 86 Stat. 103 (1972).

190. 401 U.S. at 431.

191. *Id.* at 432.

192. *Id.* It has been suggested that this case not only led to a shifting of the burden of proof but raised the standard of proof to more than mere rationality. The phrase "manifest relationship" was seen as a "new equal protection test more stringent than rational relationship but less stringent than compelling state interest." *Cultural Bias*, *supra* note 8, at 1062. For a later case restating the principles in *Griggs*, see *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975).

The court in *Riles I* found the reasoning of the jury, school, and employment cases to be persuasive.¹⁹³ Underlying the burden shift in those decisions, it said, were three principles easily analogized to the facts in *Riles I*. First, there was a "strong judicial and constitutional policy against racial discrimination."¹⁹⁴ Although the school system was not classifying students explicitly on the basis of race, in de facto discrimination cases there was "this same distrust of laws which harm blacks as an identifiable class."¹⁹⁵ Second, even when there was no intent to discriminate, there was the "existence of a positive duty to avoid racial imbalance"¹⁹⁶ on the part of school officials. Finally, disproportionality in EMR placement could not have occurred without discrimination, given the presumption that intelligence was randomly distributed among all races:

Since it does not seem to be disputed that the qualification for placement in regular classes is the innate ability to learn at the pace at which those classes proceed . . . such random distribution can be expected if there is in turn a random distribution of these learning abilities among members of all races.¹⁹⁷

Concluding its lengthy analysis, the court decided it would shift the burden of proof to the school system if the plaintiffs could demonstrate that racial imbalance existed in the composition of EMR classes and that the challenged intelligence tests were the

193. The Supreme Court has since held, however, that the "discriminatory effect" analysis used in these cases applies only to actions brought under the Civil Rights Act and not to those brought under the Constitution. *Washington v. Davis*, 426 U.S. 229 (1976). See notes 215 to 222 and accompanying text *infra*.

194. 343 F. Supp. at 1309.

195. *Id.* The court found support in *Hobson v. Hansen*, 269 F. Supp. 401 (D.D.C. 1967), for this justification. Further, the court noted that if San Francisco was indeed making explicitly racial classifications, it would have a "near impossible burden to sustain." 343 F. Supp. at 1309.

196. *Id.* at 1310. While the court did not refer to it, there was also support for this assertion in *Swann v. Board of Educ.*, 402 U.S. 1, 16 (1971), in which the Supreme Court held that when a school system failed to achieve desegregation, courts had the "broad power to fashion a remedy that will assure a unitary school system."

197. 343 F. Supp. at 1310. In fact, San Francisco had explicitly eschewed any theory of genetic inferiority. The school system did suggest that since black parents tended to be poor, and that poor pregnant women were likely to suffer from dietary deficiencies, it was plausible that a disproportionate number of black children would suffer from retarded brain development. The court, having before it no evidence to support these hypotheses, rejected them and concluded that "there can be no basis for assuming otherwise than that the ability to learn is randomly spread about the population." *Id.* at 1311.

primary determinant for placement in those classes.¹⁹⁸ Satisfied that the data as to imbalance were undisputed and unequivocal,¹⁹⁹ the court proceeded to determine whether the primary basis for EMR placement was the I.Q. test.

California's Education Code did, in fact, require a comprehensive evaluation.²⁰⁰ Students could not be placed in a program for the retarded unless their intelligence test scores were two standard deviations below the mean and a school psychologist had assessed their developmental history, cultural background, school achievement, and adaptive behavior.²⁰¹ Nevertheless, the court concluded that the school district placed emphasis on I.Q. scores; they "loom[ed] as a most important consideration in making assignments to EMR classes."²⁰² From California statutes prohibiting placement in programs for the retarded unless the I.Q. score was substantiated by other data, the court inferred that it was the score that served as the major placement determinant.

The plaintiffs, having satisfied both elements of the court's test, successfully shifted the burden to the defendant to justify its use of I.Q. tests and to "demonstrate the rational connection between the tests and the purpose for which they are allegedly used."²⁰³ The

198. *Id.* One commentator has suggested another reason for shifting the burden that could have been, but was not, relied on by the court. It has been held that where facts lie almost exclusively within the knowledge of one of the parties, the burden is on that party to support or rebut the existence of that fact. Thus, in this case, if the plaintiffs could show racial disparities in EMR placement, the information necessary to account for the discrepancy was peculiarly within the means of the defendant school system. Sorgen, *supra* note 8, at 1156.

199. "[S]tatistics often tell much, and courts listen." Alabama v. United States, 304 F.2d 583, 586 (5th Cir.), *aff'd per curiam*, 371 U.S. 37 (1962). *Accord*, Armstead v. Starkville Mun. Separate School Dist., 461 F.2d 276 (5th Cir. 1972).

200. CAL. EDUC. CODE § 6902.085 (West 1975). In 1976 and 1977 California revised and recodified its education laws. The statute controlling placement of children in classes for the retarded now appears in substantially the same form as the 1975 version at CAL. EDUC. CODE § 56506 (West 1978).

201. CAL. EDUC. CODE § 56506. The statute recognized that at the time no norms had been established and validated for adaptive behavior. It thus permitted home visits and interviews to act as a substitute. Since then, two adaptive behavior scales have been published. N. LAMBERT, M. WINDMILLER, L. COLE & R. FIGUEROA, MANUAL FOR AAMD ADAPTIVE BEHAVIOR SCALE — PUBLIC SCHOOL VERSION (1974 Version), *discussed in* J. SALVIA & J. YSELDYKE, *supra* note 4, at 371-73; J. MERCER & J. LEWIS, SYSTEM OF MULTICULTURAL PLURALISTIC ASSESSMENT (1978) (available from the Psychological Corporation, 757 Third Avenue, New York, N.Y. 10017).

202. 343 F. Supp. at 1313.

203. *Id.* at 1311. The burden of proof issue is of more than theoretical interest in a case such as this because that determination may decide the ultimate outcome of the case:

Because of the unsophisticated nature of present testing technology, it is extremely difficult for either side to clearly show that the test is valid or

school system candidly agreed that the tests were racially and culturally biased but justified their continued use on the ground that, in the absence of suitable alternatives, they were the best means available for the purpose of classifying students as retarded.²⁰⁴ The court stated: "[T]he absence of any rational means of identifying children in need of such treatment can hardly render acceptable an otherwise concededly irrational means, such as the I.Q. test as it is presently administered to black students."²⁰⁵ It held that the defendants failed to sustain "their burden of demonstrating that I.Q. tests are rationally related to the purpose of segregating students according to their ability to learn in regular classes, at least insofar as those tests are applied to black students."²⁰⁶ The school system's practices were adjudged to violate the equal protection clause. The court enjoined any future placement of black children in EMR

invalid. Therefore, the party who has the burden of proving the validity of such tests is at a distinct disadvantage. If the burden of proof did not shift, it would be as difficult for the plaintiff to show that the test is invalid as it is for the defendant to show that it is valid This shift in the burden of proof may be more important in the final analysis than whether the strict scrutiny standard or the rational relationship standard is applied.

Equal Protection, *supra* note 8, at 911.

Burden shifting in the context of classifications racially neutral on their face has been criticized. "[I]t is peculiarly inappropriate to make important social policy judgments, involving massive institutional changes, turn on the legal formalism of the burden of proof." Yudof, *supra* note 7, at 450. Another commentator believes that the court erred in not requiring the plaintiffs to show the presence of a third element before it shifted the burden, "[T]hat the use of the tests for black and poor children was not rationally related to any other legitimate educational purpose [than only intellectual capacity]," i.e., separating fast learners from slow learners. *Segregation*, *supra* note 8, at 1233. Furthermore, the commentator questioned whether courts should shift the burden under circumstances in which only disproportionate impact is shown. *Id.* See *Jefferson v. Hackney*, 406 U.S. 535, 538 (1972). There, the Supreme Court refused to shift the burden and applied traditional equal protection doctrine to a state's administration of a welfare program that heavily and negatively affected black and Mexican-American families. The Court criticized the data merely as a "naked statistical argument." But see Kirp, *Schools as Sorters*, *supra* note 8, at 759. ("[S]ince classification decisions do affect a student's educational career, a demonstration that the aptitude tests which are employed can accurately predict school performance for different types of students is an appropriate legal burden for the school to bear.").

204. San Francisco also sought to defend the disproportionality by claiming that the racial imbalance was the result of white parents more frequently than blacks placing their retarded children in private schools and of districting practices that in the past had led to the locating of EMR classes in predominantly black schools. Both contentions were rejected as lacking any substantiation in the record. 343 F. Supp. at 1313.

205. *Id.*

206. *Id.* at 1314.

classes on the basis of criteria that relied primarily on the results of intelligence tests and led to racial imbalance in such classes.²⁰⁷

Carefully wording its injunction, the court condemned only the existing method of testing black children. Despite plaintiffs' request, the court refused to order reassignment of black children already in EMR classes and upheld the practice of segregating such students.²⁰⁸ The court of appeals had no difficulty in affirming the trial court's opinion, holding that "the carefully limited relief granted [was] justified by the 'peculiar facts' of this case."²⁰⁹

Riles I may represent a modification of traditional equal protection doctrine. The court's language indicated that the defendant could meet its burden merely by showing a minimally reasonable relationship between the practice of classification and the goal of placement. However, despite the use of this conventional language, the court may have required more than simple rationality. Use of the "best means available" would clearly meet the rational basis test given a permissible end, but the court, in rejecting that defense, seemed to demand a showing of a greater congruence between the means and the end.²¹⁰ The question raised by the court,

207. For an empirically-based defense of the part school psychologists played in the evaluation of these children, see Meyers, Macmillan & Yoshida, *Validity of Psychologists' Identification of EMR Students in the Perspective of the California Decertification Experience*, 16 J. SCH. PSYCH. 3 (1978).

208. This was contrary to what would later be a strong mandate to place handicapped children in regular classrooms to the extent possible. See Education for All Handicapped Children Act, Pub. L. No. 94-142, 20 U.S.C. § 1412, 89 Stat. 781 (1976). That Act's implementing regulation, 45 C.F.R. § 121a.550 (1977), states: "[T]o the maximum extent appropriate, handicapped children . . . are [to be] educated with children who are not handicapped. . . ." From one perspective to the extent that *Riles I* preserved EMR classes, the case was a retreat from *Hobson* in which the EMR track itself was abolished. But on this point the court reversed itself in *Riles II*. See text accompanying notes 279 to 351 *infra*.

209. 502 F.2d 963, 965 (9th Cir. 1974). After the Ninth Circuit's decision was issued, the plaintiffs moved to modify the class and the terms of the preliminary injunction. The court agreed and filed an order on December 13, 1974, expanding the class to include "all Black California school children who have been or may in the future be classified as mentally retarded on the basis of I.Q. tests." *Larry P. v. Riles*, No. C-71-2270 R.F.P., slip op. at 6 (N.D. Cal. Oct. 16, 1979). The State of California decided to go even further. In 1975 it issued a memorandum and accompanying resolution stating that until further notice none of the I.Q. tests then on its approved list, which included, among others, the WISC and Stanford-Binet, could be used to place any child, regardless of race, in EMR classes. Memorandum of the Cal. Dep't of Educ., Feb. 5, 1975.

210. See Kirp, *Schools as Sorters*, *supra* note 8, at 769-70; *Cultural Bias*, *supra* note 8, at 1062-63; *Intelligence Classifications*, *supra* note 8, at 670-71; see generally Gunther, *Foreword: In Search of Evolving Doctrines on a Changing Court: A Model for a Newer Equal Protection*, 86 HARV. L. REV. 1 (1972).

translated into psychologists' terms, remains: what level of validity will meet a "substantial congruence" test? "It has not yet been determined what correlation coefficient will satisfy the various standards of review that have been applied in the [school] testing cases."²¹¹ Validity coefficients that psychologists find acceptable may not pass constitutional muster. One court has ruled that "when a program talks about labeling someone as a particular type and such a label could remain with him for the remainder of his life, the margin of error must be almost nil."²¹² "Nil" implies almost nearly perfect coefficients. Few, if any, psychometric instruments yield reliability, much less validity, coefficients above .95.

Riles I, concerned only with the propriety of a preliminary injunction, raised many legal and psychometric questions that were not answered until the trial on the merits five years later. The trial lasted more than six months, produced 10,000 pages of testimony,²¹³ and was not concluded until two years after it had begun. The outcome in *Riles II* was affected by two decisions of the Supreme Court defining the constitutional standard for proving discrimination in equal protection cases and by the passage of legislation concerning the civil rights of handicapped and retarded persons. In light of their importance for *Riles II*, the decisions and legislation are examined below.

Subsequent Supreme Court Decisions and Their Effect On Riles II

Three elements in *Riles I* assumed importance as a result of subsequent Supreme Court decisions. First, the plaintiffs claimed that the defendants had infringed their constitutional rights, not

211. *Equal Protection*, *supra* note 8, at 905 n.28. Concerning employment discrimination cases brought under Title VII, the situation is somewhat less confused because of the variety of statutory, regulatory, and professional standards available to provide evidence of job relatedness. *E.g.*, 29 C.F.R. § 1607 (1978) (EEOC Guidelines on Employee Selection Procedures).

212. *Merriken v. Cressman*, 364 F. Supp. 913, 920 (E.D. Pa. 1973). In one case teachers challenging the use of GRE scores for hiring purposes made a *prima facie* showing that the test had a disproportionately negative impact on black applicants. The court ruled that the school had failed to meet the rational relation test, and stated that "the GRE score requirement was not a reliable or valid measure for choosing good teachers." *Armstead v. Starkville Mun. Separate School Dist.*, 461 F.2d 276, 280 (5th Cir. 1972). The court's finding was based on the fact that the cutoff score would eliminate *some* good teachers. Apparently, even a few false negatives could serve to render a test arbitrary in the court's view.

213. Brief excerpts of the testimony were published in 8 APA MONITOR 4 (1977) & 9 APA MONITOR 15 (1978).

that they had violated a federal or state statute.²¹⁴ Second, the court ruled that injury resulted only from an intelligence classification that had a discriminatory effect, not from an intent to discriminate. Third, the court relied heavily on *Griggs v. Duke Power Co.*²¹⁵ to support its decision that the defendant had the burden of persuasion concerning the reasonableness of its testing practices.

The Supreme Court's 1976 decision in *Washington v. Davis*²¹⁶ had significant impact on both the decision in *Riles II* and on future testing litigation brought on grounds similar to those in that case.²¹⁷ *Davis* arose when two black applicants for positions as police officers intervened in a suit against the District of Columbia and its police department and the United States Civil Service Commission, contending that the written personnel test used by the police department excluded a disproportionately high number of black

214. In 1977, however, the court granted plaintiffs' motion to amend their complaint to include several statutory bases for their claims: Title VI of the Civil Rights Act of 1964, 42 U.S.C. § 2000d (1974); § 504 of the Rehabilitation Act of 1973, 29 U.S.C. § 794 (1974); The Education for All Handicapped Children Act, 20 U.S.C. § 1401 (1970); and several sections of the California Education Code.

215. 401 U.S. 424 (1971).

216. 426 U.S. 229 (1976).

217. There were other less significant testing cases between 1972 and 1976. In *Copeland v. School Bd.*, 464 F.2d 932 (4th Cir. 1972), plaintiffs complained that a disproportionate number of black children were assigned to classes for the retarded and learning disabled on the basis of unnamed psychological tests. The court considered special education programs, even though they isolated students, to be sound educational practices as long as placement occurred in a nondiscriminatory manner. The court found the record bare as to the reliability, validity, and nondiscriminatory attributes of the examinations used, and remanded the case to the trial court for determinations on these matters. A similar complaint was lodged by black and Mexican-American children placed in special schools for the retarded in a district that had previously operated a system segregated by law. *Arvizu v. Waco Ind. School Dist.*, 373 F. Supp. 1264 (W.D. Tex. 1973). The court ordered the school to review its testing practices to assure that assignments were made nondiscriminatorily and that children in EMR classes whose primary language was not English would be tested in their native language. In *Morales v. Shannon*, 366 F. Supp. 813 (W.D. Tex. 1973), Mexican-American children disproportionately placed in classes for the learning disabled in another Texas school system claimed a denial of equal educational opportunity in violation of the fourteenth amendment. The court, finding no evidence of discriminatory intent and determining that placement was made on the basis of objective data (not described), found no constitutional violation. Despite a finding of evidence of past discrimination and without inquiring into the validity of the intelligence and achievement tests used for placement in both EMR and programs for the gifted, a Wisconsin district court found insufficient evidence to prove that black children were inappropriately placed in the former or refused placement in the latter. *Amos v. Board of School Directors*, 408 F. Supp. 765 (E.D. Wis. 1976). The paucity of litigation generally in the mid-1970's may be explained by the passage of federal legislation providing nondiscriminatory evaluation for handicapped children. See notes 235 to 272 and accompanying text *infra*.

applicants.²¹⁸ The court of appeals applied the legal standards applicable to Title VII that had been developed in *Griggs* to resolve the intervenors' constitutional argument that the use of the test invidiously discriminated against blacks and hence denied them due process, including equal protection, guaranteed by the fifth amendment. In a decision that surprised and angered many civil rights advocates, the Supreme Court reversed the appellate court's decision, rejecting the contention that the *constitutional* standard for adjudicating claims of racial discrimination was identical to the *statutory* standard under Title VII: "[O]ur cases have not embraced the proposition that a law or other official act, without regard to whether it reflects a racially discriminatory purpose, is unconstitutional *solely* because it has a racially disproportionate impact."²¹⁹

The Court declined to use the more rigorous standard of Title VII in applying the fifth and fourteenth amendments.²²⁰ The probing judicial review under the Civil Rights Act, and its consequent lessened deference to the "seemingly reasonable acts of administrators," was inappropriate "under the Constitution where special racial impact, without discriminatory purpose, is claimed."²²¹ The Court was concerned that many statutes could be invalidated under a constitutional rule providing that a statute that was designed to serve neutral ends but which burdened one race more than another would be upheld only if there were a compelling justification.²²²

According to the Court, even under the rule established in *Davis*, a statute or practice need not reveal explicitly an intent to

218. The action in which intervention was permitted was brought by black police officers challenging the promotional policies of the department as racially discriminatory. The intervenors claimed that the test administered was not related to job performance and had a discriminatory impact in screening out black candidates. Ruling on class motions for summary judgment on the recruiting aspect of the case, the district court found that a showing sufficient to shift the burden to the defendants had been made. It concluded, however, that plaintiffs were not entitled to relief. It pointed to the department's active recruitment of blacks and the fact that 44% of recent recruits and of officers were black, a percentage equal to the percentage of blacks in the department's recruiting area. The court rejected the assertion that the test was culturally slanted and found it reasonably related to requirements of the training program. See 426 U.S. at 235-36. The court of appeals reversed the summary judgment in favor of the defendants entered by the district court.

219. *Id.* at 239.

220. The fourteenth amendment applies only to action by the states. Because this case arose in the District of Columbia, it was the fifth amendment that was applicable. However, the Court made explicit that its decision embraced constitutional challenges to alleged discrimination under both the fifth and fourteenth amendments. *Id.* at 248-49.

221. *Id.* at 247.

222. *Id.* at 248.

discriminate to be challengeable nor would a law's disproportionate impact be irrelevant.²²³ Under the Constitution invidiously discriminatory purposes may be inferred from the totality of relevant facts: evidence of systematic exclusion of one race; unequal application of the law; a demonstration that the disproportionality is difficult to explain on nonracial grounds;²²⁴ or a showing that the natural and foreseeable consequence of the state's action would have a discriminatory impact.²²⁵ Nevertheless, although disproportionate impact may be relevant, "it is not the sole touchstone of an invidious racial discrimination forbidden by the Constitution. Standing alone, it does not trigger the rule . . . that racial classifications are to be subjected to the strictest scrutiny and are justifiable only by the weightiest of considerations."²²⁶

In the second significant post-*Riles I* decision, *San Antonio Independent School District v. Rodriguez*,²²⁷ the Supreme Court held that education is not a fundamental right for purposes of equal protection analysis. Plaintiffs asserted that Texas' school financing procedures that created disparate inter-district expenditures violated the equal protection clause. More particularly, they argued that since wealth classifications were suspect, and education a fundamental right, the Supreme Court should use the compelling state interest test to scrutinize the financing system. The Court rejected this argument and, applying only the rational basis test, upheld the state's practices.²²⁸ It refused to recognize wealth as a suspect classification and held that education was not a fundamental right. Because inequitable funding did not totally deprive children of educational opportunities, "no charge fairly could be made that the system fails to provide each child with an opportunity to acquire the basic minimal skills necessary for the enjoyment of the rights of speech and of full participation in the political process."²²⁹

The Supreme Court's decisions in *Rodriguez* and *Davis* severely undercut the chances of the plaintiffs in *Riles* gaining eventual victory on the merits. Given that education could no longer be considered a fundamental right and that when a constitutional

223. *Id.* at 239.

224. *Id.* at 239-42.

225. See *Columbus Bd. of Educ. v. Penick*, 99 S. Ct. 2941 (1979); *Dayton Bd. of Educ. v. Brinkman*, 99 S. Ct. 2971 (1979) (*Dayton II*); *Dayton Bd. of Educ. v. Brinkman*, 433 U.S. 406 (1977) (*Dayton I*).

226. *Id.* at 237.

227. 411 U.S. 1 (1973).

228. *Id.* at 37.

229. *Id.*

injury was alleged minority plaintiffs must prove intent to discriminate,²³⁰ it was unlikely that the *Riles* court could employ the same reasoning that was persuasive at the preliminary injunction stage. Plaintiffs could not hope to convince the court to employ the compelling state interest test in the absence of infringement of a fundamental right or a classification based on race. Further, proving discrimination was more difficult when intent rather than effect was at issue.²³¹

The Supreme Court's decisions since 1973 have made it difficult to challenge testing procedures on constitutional grounds. School systems have a significantly greater chance to prevail when their practices are attacked under the fourteenth amendment than when they are challenged under state or federal statutes.²³² Thus, it is significant that while the Court was restricting the reach of the equal protection clause, Congress enacted a series of laws that continue to have considerable impact on the practice of psychological assessment in the public schools.

Federal Legislation

Courts develop rules of conduct in piecemeal fashion and only after litigants have presented legally cognizable issues. Rulemaking bodies such as legislatures and government agencies, on the other hand, need not wait for complaining litigants. When, among other reasons, problems need a broader solution than courts can provide,

230. In cases in which proof of a constitutional violation requires evidence of a discriminatory purpose, the burden of establishing the impermissible purpose is generally placed on the plaintiff. *Village of Arlington Heights v. Metropolitan Hous. Corp.*, 429 U.S. 252, 270-71 & n.21 (1977).

231. For suggestions on how intent may be proved, see *Dayton Bd. of Educ. v. Brinkman*, 99 S. Ct. 2971 (1979) (*Dayton II*); *Columbus Bd. of Educ. v. Penick*, 99 S. Ct. 2991 (1979); *Dayton Bd. of Educ. v. Brinkman*, 433 U.S. 406, 420 (1977) (*Dayton I*); *Village of Arlington Heights v. Metropolitan Hous. Corp.*, 429 U.S. 252, 266-68 (1977); *Lora v. Board of Educ.*, 456 F. Supp. 1211, 1282-84 (E.D.N.Y. 1978). See also Comment, *Proof of Racially Discriminatory Purpose Under the Equal Protection Clause: Washington v. Davis, Arlington Heights, Mt. Healthy, and Williamsburgh*, 12 HARV. C.R.-C.L. L. REV. 725 (1977); Comment, *Reading the Mind of the School Board: Segregative Intent and the De Facto/De Jure Distinction*, 86 YALE L.J. 317 (1976).

232. Perhaps the clearest example of this is *Lau v. Nichols*, 414 U.S. 563 (1974), in which Chinese parents alleged that the San Francisco school system denied their children equal protection in failing to provide bilingual instruction. Though the parents brought the claim under the fourteenth amendment, the Supreme Court chose to rely solely on the Civil Rights Act of 1964, 42 U.S.C. § 2000d (1970), to find discrimination on the basis of national origin. Under the act, intent to discriminate need not be shown. Finding that "[d]iscrimination is barred which has that effect even though no purposeful design is present," 414 U.S. at 568, the Court concluded that the school system denied the plaintiffs a meaningful opportunity to participate in a public educational program in violation of the act. *Id.* at 567-68.

or they affect many people, lawmakers enact statutes and administrators promulgate regulations that have comprehensive effect. This process is evident with regard to special education and psychological testing. Legislation for the handicapped and those misclassified as handicapped directly affects the assessment practices of psychologists. The involvement of the federal government in the education of handicapped children began in 1966²³³ but did not become intensive until 1974, when Congress passed the Education of the Handicapped Amendments, more commonly known to educators as Public Law 93-380.²³⁴

1. *Education of the Handicapped Amendments of 1974*

Public Law 93-380 reflected the impact of the litigation of the preceding decade concerning the exclusion of handicapped children from public schools²³⁵ and the discriminatory evaluation and placement practices used against minority children. Its provisions concerning special education,²³⁶ however, are now only of historical interest as they were amended significantly by Public Law 94-142 in 1975.²³⁷ Public Law 93-380 did place school systems on notice that for the first time federal financial assistance for special education would be contingent on the development by state educational agencies of plans establishing "a goal of . . . full educational opportunity to all handicapped children"²³⁸ and a means of

233. In that year, Title VI was added to the Elementary and Secondary Education Act of 1965, Pub. L. No. 89-750, § 161, 80 Stat. 1204 (1966), *amending* Pub. L. No. 89-10, 79 Stat. 27 (1965) (establishing the Bureau of Education for the Handicapped). For a history of the federal government's development of legislation for the handicapped, see *Lora v. Board of Educ.*, 456 F. Supp. 1211, 1224-30 (E.D.N.Y. 1978); Note, *The Education for All Handicapped Children Act of 1975*, N.Y.U. J. L. REF. 110, 119-20 (1976). A parallel development occurred at the state level. By 1979, 49 states by law provided free public education to their handicapped citizens. See Connor & Connors, *Children's Rights and Mainstreaming of the Handicapped*, in *CHILDREN'S RIGHTS: CONTEMPORARY PERSPECTIVES* 67, 69 (P. Vardin & I. Brody eds. 1979).

234. Pub. L. No. 93-380, §§ 611-621, 88 Stat. 579 (1974) (codified at 20 U.S.C. §§ 1401-1461 (Supp. IV 1974)), *amending* 20 U.S.C. §§ 1401-1461 (1970).

235. *E.g.*, *Mills v. Board of Educ.*, 348 F. Supp. 866 (D.D.C. 1972); *Pennsylvania Ass'n for Retarded Children v. Pennsylvania*, 343 F. Supp. 279 (E.D. Pa. 1972).

236. Public Law 93-380 was an omnibus bill containing seven titles, of which only Title VI(B) affected special education. Other titles amended the Elementary and Secondary Education Act, reaffirmed policies concerning equal educational opportunity, and consolidated certain other education programs. Of particular relevance is a portion of Title V that established the Family Education Rights and Privacy Act, discussed in the text accompanying notes 426 to 434 *infra*.

237. Education for All Handicapped Children Act, Pub. L. No. 94-142, 89 Stat. 773 (1975). See text accompanying notes 242 to 266 *infra*.

238. 20 U.S.C. § 1413(12)(A) (Supp. IV 1974).

"insuring that handicapped children and their parents or guardians are guaranteed procedural safeguards in decisions regarding identification, evaluation, and educational placement. . . ."239 One such safeguard was that parents were to be given the right to "examine all relevant records with respect to . . . classification or educational placement, and obtain an independent educational evaluation of the child."240 To prevent misclassification of minority children, school systems were to develop methods for insuring that any assessment devices used "for the purposes of classification and placement of handicapped children will be selected and administered so as not to be racially or culturally discriminatory."241 These broad mandates were particularized in Public Law 94-142, discussed below.

2. *Education for All Handicapped Children Act*²⁴²

Public Law 94-142, the Education for All Handicapped Children Act, is a grant-giving statute providing financial support to state and local education agencies for special education services if they meet certain detailed eligibility requirements. The purpose of the Act is to

assure that all handicapped children have available to them . . . a free appropriate public education which emphasizes special education and related services designed to meet their unique needs, to assure that the rights of handicapped children and their parents or guardians are protected, to assist States and localities to provide for the education of all handicapped children, and to assess and assure the effectiveness of efforts to educate handicapped children.²⁴³

239. *Id.* § 1413(13).

240. *Id.* § 1413(13)(A)(ii).

241. *Id.* § 1413(13)(C).

242. Pub. L. No. 94-142, 89 Stat. 773 (1975) (codified at 20 U.S.C. §§ 1401-1461 (1978)).

243. 20 U.S.C. § 1401(c) (1978). The means for achieving these purposes are set out in the rest of the statute and the implementing regulations drafted by the Office of Education in the Department of Health, Education and Welfare. 45 C.F.R. § 121a.1-.754 (1977). Because of the difficulty in defining and assessing children suspected of learning disability, see 41 Fed. Reg. 52, 404 (1976), Congress ordered HEW to develop procedures for identifying and evaluating such children. These regulations were not published with the main body of regulations and appear at 42 Fed. Reg. 65,082 (1977). They require a multidisciplinary evaluation including observation of the child's academic performance in the regular classroom setting. 45 C.F.R. § 121a.540, .542. See text accompanying notes 391 to 395 *infra*.

Public Law 94-142 places constraints on the administration of psychological tests and provides standards for evaluating children suspected of having handicaps. Under the implementing regulations a school system must obtain written parental consent before conducting a "preplacement evaluation."²⁴⁴ The consent form must state the purpose of the evaluation and describe each evaluation procedure or test the school proposes to use.²⁴⁵ Other kinds of assessment, such as large-scale screening to identify those children who might need more intensive individual psychological evaluation, fall outside the definition of preplacement testing, although parents do have the right to be notified that such screening will occur.²⁴⁶ It is likely that observation to assess classroom atmosphere and interactions, or for screening purposes, also would not require consent, because involvement of children in these processes is relatively insignificant. Even covert observation of public behavior does not sufficiently infringe privacy rights as to require informed consent.²⁴⁷ Further, because the possibility of labeling and placement is remote, the need for strict procedural protections, like consent, is minimal. However, when a particular child becomes the focus of an assessment, the purpose of which is to determine the need for special education placement, then parental consent seems required for all procedures, including testing, interviewing, and observation.²⁴⁸

In an earlier draft the regulations gave parents the power to veto preplacement evaluations.²⁴⁹ That proposal failed to take into account the possibility of conflicting interests between a parent and a child in need of special education. It would have been possible under the proposed rules for parents without valid reasons to deny children access to psychological services and remedial intervention

244. 45 C.F.R. § 121a.504 (1977). Under the statute itself, only notice, not consent, was required. 20 U.S.C. § 1415(b)(1)(c) (1978).

245. 45 C.F.R. § 121a.505(a)(3) (1977). Methods for implementing this provision have been suggested by Bersoff, *The Implementation of Procedural Safeguards Under Pub. L. 94-142*, in DEVELOPING CRITERIA FOR THE EVALUATION OF DUE PROCESS PROCEDURAL SAFEGUARDS PROVISIONS 65, 98-100 (Bureau of Educ. for the Handicapped ed. 1978).

246. 45 C.F.R. § 121a.504(a)(1) (1977).

247. Bersoff, *Legal and Ethical Concerns in Research*, in RESEARCH FOR COUNSELORS 386 (L. Goldman ed. 1978).

248. While the regulations require comprehensive reevaluations of handicapped children every three years, 45 C.F.R. § 121a.534(b) (1977), they do not mandate consent for this purpose. Such consent has been recommended, however. Bersoff, *supra* note 245, at 92.

249. 41 Fed. Reg. 56,990 (1976) (proposed regulation § 121a.404(b)).

by special educators.²⁵⁰ The final regulations are preferable because they permit schools to act as children's advocates by providing alternative mechanisms to challenge parental refusal to consent.²⁵¹

It might have been more efficient and economical to develop less rigorous requirements to deal with parents' refusal to consent to preplacement evaluations because such requirements place constraints on school systems and create delay in providing services. Actions that precede placement, such as the administration of psychological tests, must await the outcome of hearings called to determine if the right of parents to act as care givers and decisionmakers for their children should be supplanted by the school. However, the rationale behind the regulations is sound. By requiring parental consent, allowing school officials to challenge refusal to consent, and developing a forum in which both sides will be heard by a neutral adjudicator, the regulations serve the interests

250. The regulations require a preplacement evaluation before any programming decisions can be made. 45 C.F.R. § 121a.531 (1977). Further, parents must consent to initial admission to special education as well as to the initial evaluation. *Id.* § 121a.504(b)(ii).

251. Depending on state statutes, the regulations establish two means that education agencies can use to initiate impartial adjudications to override parental vetoes, state neglect laws, or hearing procedures developed in Public Law 94-142. Some states already have provisions concerning parental consent. For example, Maryland's State Department of Education Bylaws declare: "Parents . . . have the right of prior informed consent regarding their child's psychological evaluation, . . . special education programming and placement in accordance with procedures established by each local education agency." Code of Bylaws, Md. State Bd. of Educ. § 13.04.01.03. Under this law, parents apparently do possess veto power. However, the school system may have recourse to state parental neglect statutes. A school administrator in some states may file a neglect petition when parents refuse to consent to evaluation or placement. Maryland law defines a neglected child as one "whose parent . . . legally responsible for his care does not adequately supply him with . . . education" Md. CTS. & JUD. PROC. CODE ANN. § 3-801 (1974). In these instances, anyone having knowledge of facts regarding neglect may file a complaint with the juvenile, or other appropriate, court. Such action will trigger an inquiry into possible neglect and can result in a hearing in which the court may appoint a guardian of the child for the limited purpose of consenting to evaluation or placement. The procedure, while lengthy, does allow school systems to provide needed services in those cases in which parents refuse to consent, and state laws, bylaws, or regulations require consent.

Where no state law exists requiring consent, a refusal to consent may be overridden by following hearing and appeal procedures delineated in other sections of the regulations implementing Public Law 94-142. 45 C.F.R. § 121a.503(c)(2)(i), .506-.511 (1977). These provisions require that a hearing be conducted before an impartial adjudicator. Both parents and the school may be represented by attorneys and have the opportunity to present evidence and confront, cross-examine, and compel the attendance of witnesses. After the hearing is completed, the hearing officer must issue a written decision based on the facts. Either party may appeal the decision to a state hearing panel and ultimately to the courts.

of all the parties: the parents, whose constitutional right to direct their children's upbringing is protected;²⁵² the school, which can carry out its statutory duty to provide an appropriate education for handicapped children; and, most important, the children, whose future is determined by the actions of the adults around them.

Public Law 94-142 reaffirmed Public Law 93-380's mandate concerning nondiscriminatory evaluation. To be eligible for funding, states must establish procedures to insure that testing and evaluation devices are neither culturally nor racially discriminatory.²⁵³ The implementing regulations clarify this requirement:

Tests and other evaluation materials:

- (1) Are provided and administered in the child's native language or other mode of communication . . . ;
- (2) Have been validated for the specific purpose for which they are used; and
- (3) Are administered by trained personnel in conformance with the instructions provided by their producer²⁵⁴

The most ambiguous of these provisions is section (2). The regulations require test validation but not test validity. Even if one infers that both are necessary, the regulations do not specify to what level of validity a test must conform. There are few judicial or statutory guidelines concerning standards of validity in school testing,²⁵⁵ and there is little agreement on the general concept of nondiscriminatory assessment.²⁵⁶

252. *E.g.*, *Parham v. J.R. & J.L.*, 99 S. Ct. 2493 (1979); *Wisconsin v. Yoder*, 406 U.S. 205 (1972).

253. 20 U.S.C. § 1414(5)(c) (1978). Similar assurances are required elsewhere. The Office for Civil Rights has interpreted Title VI of the Civil Rights Act of 1964, 42 U.S.C. § 2000d (1974), to compel all school systems receiving federal financial assistance to develop procedures preventing disproportionate overinclusion or underinclusion of children of any race, color, national origin, or sex in all special programs. Office for Civil Rights, Memorandum for Chief State School Officers and Local School District Superintendents (August 1975). For a fuller discussion of federal regulations in this context, see *Oakland & Laosa, supra* note 8, at 28-37.

254. 45 C.F.R. § 121a.532(a) (1977).

255. See text accompanying notes 211 & 212 *supra*. The Senate Report summarizing and explaining the provisions of Public Law 94-142 merely reiterates the basic requirement of nondiscriminatory evaluation without providing guidance as to the section's meaning. S. Rep. No. 168, 94th Cong., 1st Sess. 47 (1975). *Larry P. v. Riles*, No. C-71-2270 R.F.P. (N.D. Cal. Oct. 16, 1979), *appeal docketed*, No. 80-4027 (9th Cir. Jan. 17, 1980), provides a definition of validity. See text accompanying notes 296 to 298 *infra*.

256. See text accompanying notes 115 to 149 *supra*.

Other provisions also affect psychological and educational assessment. Children with sensory, manual, or speaking impairments must be given tests that measure genuine aptitude or achievement, not simply impairments.²⁵⁷ Further, all assessment is to be comprehensive, multifaceted, and multidisciplinary.²⁵⁸ Evaluations for placement must be conducted by persons from the fields of education, medicine, and psychology, who assess children "in all areas related to the suspected disability, including, where appropriate, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities."²⁵⁹ Tests that are "merely . . . designed to provide a single general intelligence quotient"²⁶⁰ may not be used, nor can a single procedure be "used as the sole criterion for determining an appropriate educational program for a child."²⁶¹ As salutary as this latter protection is, it is relatively meaningless in a situation such as *Riles* where the issue was the use of I.Q. scores as the primary, not the sole, determinant for placement. Few, if any, schools use only an intelligence test for this purpose.

To insure that all provisions are given effect, both the statute²⁶² and the regulations²⁶³ enable parents "to present complaints with respect to any matter relating to the identification, evaluation, or educational placement"²⁶⁴ of their children. The complaints are presented in an impartial administrative hearing in which parents have the right to compel the attendance of, and to cross-examine, witnesses involved in the assessment and programming decisions.²⁶⁵ Psychologists called to testify may have their credentials, and their methods, interpretations, and recommendations subjected to the closest scrutiny.²⁶⁶

257. 45 C.F.R. § 121a.532(c) (1977).

258. *Id.* § 121a.532(a).

259. *Id.* § 121a.532(e) & (f).

260. *Id.* § 121a.532(b).

261. *Id.* § 121a.532(d). In making placement decisions, the school is required to "draw upon information from a variety of sources, including aptitude and achievement tests, teacher recommendations, physical condition, social and cultural background, and adaptive behavior." *Id.* § 121a.533(a)(1).

262. 20 U.S.C. § 1415(e)(2) (1978).

263. 45 C.F.R. § 121a.500-.514 (1977).

264. 20 U.S.C. § 1415(b)(1)(E) (1978).

265. 45 C.F.R. § 121a.508 (1977).

266. See J. ZISKIN, *COPING WITH PSYCHIATRIC AND PSYCHOLOGICAL TESTIMONY* (1975) (instructing attorneys on how to challenge the credentials, expertise, and psychometric soundness of psychologists' evaluation devices); Comment, *The Psychologist as Expert Witness: Science in the Courtroom?*, 38 MD. L. REV. 539 (1979). See also Morse, *Crazy Behavior, Morals, and Science: An Analysis of Mental Health Law*, 51 S. CAL. L. REV. 527 (1978).

3. *Rehabilitation Act of 1973*²⁶⁷

In section 504 of the Rehabilitation Act of 1973, which was enacted to promote the education and training of handicapped persons, Congress declared:

No otherwise qualified handicapped individual in the United States . . . shall, solely by reason of his handicap, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance.²⁶⁸

This language embodies "the first Federal civil rights law protecting the rights of handicapped persons and reflects a national commitment to end discrimination on the basis of handicap."²⁶⁹ Section 504 is almost identical to Title VI of the 1964 Civil Rights Act and may be just as comprehensive.²⁷⁰ Unlike Public Law 94-142, the requirements of section 504 are not triggered solely by receipt of funds under a specific statute, but protect handicapped persons in all institutions receiving federal financial assistance. Any school system, public or private, receiving federal monies for any program or activity whatsoever is bound by its mandates.

In mid-1977, regulations implementing the broad right-granting language of section 504 were issued.²⁷¹ In addition to general principles already established under Public Law 94-142, they include rules for the evaluation of children suspected of being handicapped, which require preplacement evaluations, validated tests, and multidisciplinary comprehensive assessment.²⁷²

One possible advantage of section 504 is that it may give a private right of action to plaintiffs challenging the assessment practices of school officials. Because section 504 so resembles the 1964 Civil Rights Act, it may be assumed that the remedies available under the act would also be available under section 504. *Regents of University of California v. Bakke*²⁷³ left inconclusive whether Title

267. 29 U.S.C. §§ 701-794 (1975).

268. Pub. L. No. 93-112, § 504, 87 Stat. 394 (1973), *as amended by* Pub. L. No. 95-602, § 119, 92 Stat. 2982 (1978) (codified at 29 U.S.C. § 794 (1979)).

269. 42 Fed. Reg. 22,676 (1977).

270. *See* note 232 *supra*.

271. 45 C.F.R. §§ 84.1-.61 (1977).

272. 45 C.F.R. § 84.35 (1978). Because they so closely track the Public Law 94-142 regulations, *see* text accompanying notes 254 to 261 *supra*, they will not be repeated here.

273. 438 U.S. 265 (1978).

VI of the Civil Rights Act establishes a private cause of action. The Court avoided the issue whether Section 504 itself afforded such a right in *Southeastern Community College v. Davis*,²⁸⁴ the first case in which it had the opportunity to construe the meaning of that provision.²⁷⁵ However, lower courts have interpreted Section 504 to enable individuals to bring private causes of action for discrimination in education,²⁷⁶ transportation,²⁷⁷ and employment.²⁷⁸

Larry P. v. Riles: *Decision on the Merits*²⁷⁹

Federal legislation protecting handicapped persons and other civil rights statutes had a significant effect on the outcome of *Riles II*. The plaintiffs twice amended their complaint to allege claims based on several federal statutes²⁸⁰ in addition to their equal

274. 99 S. Ct. 2361 (1979).

275. *Id.* at 2366 n.5. *Cf.* *Regents of Univ. of Cal. v. Bakke*, 438 U.S. 265, 420 n.27 (1978) (Stevens, J., concurring) (citing the legislative history of section 504 as support for his conclusion that a person may maintain a cause of action under Title VI).

276. *E.g.*, *Kampmeier v. Nyquist*, 553 F.2d 296 (2d Cir. 1977); *Larry P. v. Riles*, No. C-71-2270 R.F.P. (N.D. Cal. Oct. 16, 1979), *appeal docketed*, No. 80-4027 (9th Cir. Jan. 17, 1980); *Boxall v. Sequoia Union High School Dist.*, 464 F. Supp. 1104 (N.D. Cal. 1979); *Stemple v. Board of Educ.*, 464 F. Supp. 258 (D. Md. 1979), *appeal docketed*, No. 79-1208 (4th Cir. March 30, 1979); *Howard S. v. Friendswood Ind. School Dist.*, 454 F. Supp. 634 (S.D. Tex. 1978); *Barnes v. Converse College*, 436 F. Supp. 635 (D.S.C. 1977).

277. *E.g.*, *United Handicapped Fed'n v. Andre*, 558 F.2d 413 (8th Cir. 1977); *Lloyd v. Regional Transp. Auth.*, 548 F.2d 1277 (7th Cir. 1977).

278. *E.g.*, *Whitaker v. Board of Educ.*, 461 F. Supp. 99 (E.D.N.Y. 1978); *Drennon v. Philadelphia Gen. Hosp.*, 428 F. Supp. 809 (E.D. Pa. 1977); *Gurmankin v. Costanzo*, 411 F. Supp. 982 (E.D. Pa. 1976), *aff'd on other grounds*, 556 F.2d 184 (3d Cir. 1977); *cf.* *Coleman v. Darden*, 595 F.2d 533 (10th Cir. 1979) (implying private cause of action against private employers but holding no such right against federal agency).

Title V of the Rehabilitation Act was amended in November 1978, specifically permitting remedies identical to those in the 1964 Civil Rights Act. Pub. L. No. 95-602, § 120(a), 92 Stat. 2982, 2987 (1978) (to be codified at 29 U.S.C. § 794(a)(2)). One court, equating the amendment with the restrictions on remedies with that of Titles VI and VII of the 1964 Civil Rights Act, has construed the amendment to preclude a private cause of action under § 504 to redress employment discrimination unless the primary objective of federal financial assistance is to provide employment. *Trageser v. Libbie Rehab. Center, Inc.*, 590 F.2d 87 (4th Cir. 1978). The Fourth Circuit's earlier holding recognizing a private cause of action under § 504 in *Davis v. Southeastern Community College*, 574 F.2d 1158 (4th Cir. 1978), *rev'd on other grounds*, 442 U.S. 397 (1979), was distinguished in that the plaintiffs in *Davis* alleged discrimination in an educational institution and therefore the restrictive remedies in employment cases were not germane. *But see* *Hart v. County of Alameda*, 21 Empl. Prac. Dec. 233 (N.D. Cal. 1979).

279. No. C-71-2270 R.F.P. (N.D. Cal. Oct. 16, 1979), *appeal docketed*, No. 80-4027 (9th Cir. Jan. 17, 1980) (summarized at 48 U.S.L.W. 2298 (Oct. 30, 1979)).

280. *See* note 214 *supra*.

protection claim.²⁸¹ Finding these additional claims to be persuasive, the court held for the plaintiffs on both statutory and constitutional grounds, permanently enjoining the defendants "from utilizing, permitting the use of, or approving the use of any standardized intelligence tests . . . for the identification of black E.M.R. children or their placement into E.M.R. classes, without securing prior approval by this court."²⁸² Presumably because proof of discrimination under the statutes required plaintiffs to show only that the state's actions created an impermissible discriminatory impact and not an intent to discriminate, the court began its analysis by evaluating the plaintiffs' claims under Title VI of the Civil Rights Act,²⁸³ the Rehabilitation Act,²⁸⁴ and the Education for All Handicapped Children Act.²⁸⁵

In its analysis of plaintiffs' cause of action under Title VI, the court relied on *Lau v. Nichols*,²⁸⁶ in which the Supreme Court had held that English instruction for children whose primary language was not English amounted to functional exclusion from schooling. The court reasoned that placement mechanisms for EMR classes that effectively foreclosed a disproportionate number of black children from any meaningful education also violated Title VI.²⁸⁷ The defendants were unable to rebut the plaintiffs' prima facie showing of disproportionate and invidious impact of its assessment and placement practices: there was no evidence that EMR classes "actually reflected and tapped a greater incidence of mild mental retardation" in black children or that the intelligence tests used for placement in those classes "had been validated for [that] purpose."²⁸⁸

The court reserved most of its statutory analysis for plaintiffs' claims under section 504 and Public Law 94-142, focusing on the

281. Plaintiffs' amended complaint also alleged violation of the Emergency School Aid Act of 1978, 20 U.S.C. §§ 3191-3207 (1979). The court held that plaintiffs had not demonstrated the applicability of the act to California's school system and so dismissed their claims under those provisions. No. C-71-2270 R.F.P., slip op. at 125 n.75 (N.D. Cal. Oct. 16, 1979). See note 214 *supra*. The United States Department of Justice, participating as amicus curiae, asserted that the defendants' conduct also violated Public Law 94-142 and the Rehabilitation Act of 1973.

282. *Id.* at 104.

283. 42 U.S.C. § 2000d (1974).

284. 29 U.S.C. § 794 (1974).

285. 20 U.S.C. § 1401 (1974).

286. 414 U.S. 563 (1974). See notes 177 & 232 *supra*.

287. Slip op. at 58-59.

288. *Id.* at 61.

nondiscriminatory assessment provisions of these acts,²⁸⁹ particularly the requirement of the regulations that assessment instruments be "validated for the specific purpose for which they are used."²⁹⁰ The court's interpretation of these provisions was of crucial importance in its holding and the shaping of the final remedy; through it, the court broke new ground.²⁹¹ As it had in *Riles I*, the court relied on *Griggs v. Duke Power Co.*²⁹²

There, the Supreme Court held that to rebut a prima facie case of discrimination brought by employees who claim an employer's use of tests creates a disproportionate impact on minorities, an employer must show that the test has a "manifest relationship" to the position for which the test is required.²⁹³ If this is done, the burden of production shifts to the plaintiffs, who may then submit evidence that alternative selection procedures exist that would serve the employer's purposes as well without producing discriminatory effects.²⁹⁴ In *Riles II*, the court accepted the burden-shifting approach but found it impossible to translate *Griggs*' "manifest relationship" test to the educational setting:

If tests can predict that a person is going to be a poor employee, the employer can legitimately deny that person a job, but if tests suggest that a young child is probably going to be a poor student, the school cannot on that basis alone deny that child the opportunity to improve and develop the academic skills necessary to success in our society. Assignment to E.M.R. classes denies that opportunity through relegation to a markedly inferior, essentially dead-end track.²⁹⁵

As an alternative to the "manifest relationship" test, the court held that the defendant should bear the burden of proving that the tests used for placement had been validated for black children. However, it would not accept proof merely that the tests used were

289. See text accompanying notes 248 to 272 *supra*. The court also chastised the state for adopting policies contrary to the least restrictive environment, or "mainstreaming," provisions of the statutes, 20 U.S.C. § 1412 (1978), 29 U.S.C. § 794 (1974), which require that "to the maximum extent appropriate, handicapped children . . . are [to be] educated with children who are not handicapped . . .," 20 U.S.C. § 1412 (1978). Slip op. at 61.

290. 45 C.F.R. § 121a.532 (1977).

291. "There are no cases applying validation criteria to tests used for E.M.R. Placement." Slip op. at 64.

292. 401 U.S. 424 (1971).

293. *Id.* at 430.

294. *Id.* at 431.

295. Slip op. at 66.

able to predict school performance. Rather, it adopted a more stringent requirement that the tests be shown valid for selecting children unable to profit from instruction in regular classes with remedial instruction.²⁹⁶ The tests would have to identify accurately children who belonged in what the court characterized as "isolated, dead-end, stigmatizing E.M.R. programs."²⁹⁷ This kind of validation, the court found, had not been done. "[D]efendants must come forward and show that [the tests] have been validated for each minority group with which they are used. . . . This minimal burden has not been met for diagnosing the kind of mental retardation justifying E.M.R. placement."²⁹⁸

Continuing with the analysis in *Griggs*, the court found that alternative mechanisms for determining placement in EMR classes existed. Between 1975 and the resolution of the case in 1979, there had been a state-wide moratorium on the use of I.Q. tests to place children of all races in EMR programs. The state's employees testified that adequate assessments had been made during that period without I.Q. tests and that there was no evidence to suggest that misplacements had occurred.²⁹⁹ The court found, in fact, that more time and care had been taken during this period in placing children in EMR classes:

School psychologists, teachers, and others involved in the process are now making decisions based on a wide number of factors, and the evidence suggests that the results are less discriminatory than they were under the I.Q.-centered standard. Evaluations can and are taking place through, *inter alia*, more thorough assessments of the child's personal history and development, adaptive behavior both inside and outside of the school environment, and classroom performance and academic achievement.³⁰⁰

296. *Id.* at 67.

297. *Id.* at 68.

298. *Id.* at 69. The few studies that had been brought to the court's attention were not considered relevant. The court rejected validity studies correlating I.Q. scores with college grades or with other achievement tests. It would be satisfied only with research relating I.Q. scores of black children with classroom grades. The one relevant study cited yielded correlations between I.Q. scores and grades for white children of .25 and only .14 for blacks. The expert who testified about the study concluded that the Wechsler Intelligence Scale for Children had "little or no validity for predicting the scholastic performance of black or brown children." *Id.* at 71. Thus, the court concluded that "the I.Q. tests are differentially valid for black and white children. . . . Differential validity means that more errors will be made for black children than whites, and that is unacceptable." *Id.* at 71-72.

299. *Id.* at 73.

300. *Id.* at 73-74. Several psychologists have argued that systematic observation of situation-specific behavior — such as that in a classroom — may be a more valid

Nevertheless, the court warned, alternatives to I.Q. tests themselves had not been validated, and disproportionate placement, while less egregious than in the pre-1975 era, was still present. Continued use of tests would still be needed, not for the purpose of labeling children as retarded, but for "the development of curricula that respond to specific educational needs."³⁰¹ Thus, given the functional exclusion of black children under Title VI and the state's failure to meet its burdens under section 504 and Public Law 94-142, the court found that the defendant had violated those statutes.³⁰²

Because testing for EMR placement had been enjoined in *Riles I* on the basis of the plaintiffs' constitutional claims, the court considered itself obliged to determine whether the plaintiffs continued to warrant relief under the equal protection clause "where this litigation commenced."³⁰³ Given *Davis* and its progeny, the plaintiffs had a more difficult task under the fourteenth amendment than they had under their statutory claims: "The difficult question of intent [had] moved to center stage."³⁰⁴ The court believed that the problem was made more burdensome because of what it perceived to be the failure of the federal judiciary to delineate a precise formula for determining intentional discrimination under *Davis*.³⁰⁵ It wished to avoid relying solely on an objective test that would measure intent by establishing that the state's conduct had the natural and foreseeable consequences of producing a discriminatory effect.³⁰⁶

measure of performance than the artificial environment of the testing room. *E.g.*, E. MASH & L. TERDAL, BEHAVIOR THERAPY ASSESSMENT 261-352 (1976); D. PETERSON, THE CLINICAL STUDY OF SOCIAL BEHAVIOR (1968); R. THARP & R. WETZEL, BEHAVIOR MODIFICATION IN THE NATURAL ENVIRONMENT (1969); M. TILLMAN, D. BERSOFF & J. DOLLY, LEARNING TO TEACH: A DECISION-MAKING SYSTEM 262-329 (1976).

301. Slip op. at 74. This recommendation comports with the modern commentary on educational assessment. *E.g.*, J. SALVIA & J. YSSELDYKE, *supra* note 4, at 472-74; Bersoff, *supra* note 7; McClelland, *supra* note 7.

302. Slip op. at 100-01.

303. *Id.* at 75. "[T]his case began strictly as a constitutional one, and the equal protection claim has remained central." *Id.* at 55. The court initially determined that it would employ the intermediate standard of review. *Id.* at 76-84. *See* note 185 *supra* and text accompanying notes 352 to 354 *infra*.

304. Slip op. at 75.

305. *Id.* at 76.

306. *Id.* at 77. This approach was suggested by Justice Stevens' concurring opinion in *Washington v. Davis*, 426 U.S. 229, 253 (1976). Several lower federal courts had used that approach in post-*Davis* cases, *e.g.*, *Arthur v. Nyquist*, 573 F.2d 134 (2d Cir.), *cert. denied*, 439 U.S. 860 (1978); *NAACP v. Lansing Bd. of Educ.*, 559 F.2d 1042 (6th Cir.), *cert. denied*, 434 U.S. 997 (1977); *United States v. Texas Educ. Agency*, 532 F.2d 380 (5th Cir.), *vacated and remanded sub nom. Austin Ind. School Dist. v. United States*, 429 U.S. 990 (1976), 564 F.2d 162 (5th Cir. 1977), *petition for reh. denied*, 579 F.2d 910 (5th Cir. 1978), *cert. denied*, 99 S. Ct. 3106 (1979); *Lora v. Board of Educ.*, 456 F. Supp. 1211 (E.D.N.Y. 1978), but the Ninth Circuit — whose decisions bound the

The court concluded that although the plaintiffs would have to prove more than a foreseeable discriminatory impact, it would not so narrowly define discriminatory purpose to mean "an intent to harm black children."³⁰⁷ The plaintiffs would have to show an intent to segregate those children into classes for the educable retarded.³⁰⁸ This they had accomplished. The court found, after a detailed and lengthy analysis of California's education system generally and its programs for retarded children specifically, that the state had been unable to meet the educational needs of disadvantaged children for most of its history and viewed placement of blacks in EMR classes as but one aspect of this failure.³⁰⁹ However, it was the EMR program that received the brunt of the court's condemnation.³¹⁰ Relying on either the testimony of the state's own witnesses or on its documents, the court concluded that EMR classes were "designed to separate out children who are *incapable* of learning in the regular classes"³¹¹ and were not meant to provide remedial instruction so that children could learn the skills necessary for eventual return to regular instruction.³¹² Given these characteristics, the court considered "the decision to place children in these classes . . . a crucial one. Children wrongly placed in these classes are unlikely to escape as they inevitably lag farther and farther behind the children in regular classes."³¹³

The court next reviewed the process by which a disproportionate number of black children were placed in EMR classes.³¹⁴ Based on

Riles court — had not yet explicitly adopted the objective test. *Soria v. Oxnard School Dist.*, 488 F.2d 579 (9th Cir. 1973). *But see De La Cruz v. Tormey*, 582 F.2d 45 (9th Cir. 1978), *cert. denied*, 441 U.S. 965 (1979). Moreover, a few months prior to *Riles II*, the Supreme Court had warned that while "inevitability or foreseeability of consequences [permits] a strong inference that the adverse effects were desired [it was but] . . . a working tool, not a synonym for proof." *Personnel Adm'r v. Feeney*, 442 U.S. 256, 279 n.25 (1979). *See also* *Columbus Bd. of Educ. v. Penick*, 99 S. Ct. 2941 (1979); *Dayton Bd. of Educ. v. Brinkman*, 99 S. Ct. 2971 (1979).

307. Slip op. at 84.

308. *Id.*

309. *Id.* at 1.

310. Throughout the opinion the court labeled it "dead-end," "isolating," "inferior," and "stigmatizing." *Id.* at 18, 24, 66, 68, 85, 92, 101.

311. *Id.* at 16.

312. *Id.* at 17.

313. *Id.* at 18-19.

314. The court found that in the 20 school districts in California "accounting for 80 percent of the enrollment of black children, black children comprised 27.5 percent of the student population and 62 percent of the E.M.R. population." *Id.* at 23. This overrepresentation, the court concluded, could have occurred by chance less than one time in a million. *Id.* Further, even if one assumed that the incidence of mild mental retardation was 50 percent greater in black children than white, there was "still less than a one in 100,000 chance that the enrollment could be so skewed towards black children." *Id.* at 24.

the testimony of the state's witnesses it found that although California had acknowledged in 1969 that minorities were overrepresented in EMR classes, it chose for the first time in that year to order the use of specific standardized individual intelligence tests for EMR placement.³¹⁵ The "quick and unsystematic" process by which tests were selected for use³¹⁶ failed to consider "critical issues stemming from I.Q. testing,"³¹⁷ by relying on the most commonly used tests, the state "opted to perpetuate any discriminatory effects of those tests."³¹⁸ Before it would decide whether the tests were indeed discriminatory, the court determined, as it had in *Riles I*, that I.Q. tests were the primary determinant in EMR placement,³¹⁹ relying on the results of the state's own investigation.³²⁰

These initial analyses finally brought the court to the central issue — the nature of the intelligence tests themselves. Expert witnesses for both plaintiffs and defendants had agreed in two crucial respects: first, it was impossible to "truly define, much less measure, intelligence," instead, "I.Q. tests, like other ability tests, essentially measure achievement . . . ;"³²¹ and second, black children performed significantly less well on intelligence tests than did their white counterparts.³²² Only two percent of white students in California achieved I.Q. scores below 70, while fifteen percent of black students did. The court first asked why the tests had not been modified to remove this disparity in the same way that the

315. *Id.* at 26. The tests included the Stanford-Binet Intelligence Scale, the Wechsler Intelligence Scale, and the Leiter International Performance Scale. The list was further amended in 1975 to include a number of supplemental intelligence tests. *Id.* at 111 n.3, 117 n.39.

316. The list of such tests had been developed by a state department of education official who was not an expert in I.Q. testing, and who had made his selections on the basis of a survey of the tests used most frequently by California's school psychologists and on the recommendations of test publishers. *Id.* at 27.

317. *Id.*

318. *Id.* at 28. The list was approved summarily by the California Board of Education on the assumption that the selection of tests had been based on expert opinion with due consideration for the problems of cultural bias. *Id.*

319. *Id.* at 30-31.

320. The state's own investigation revealed that about one-third of EMR pupils' records contained no estimates of adaptive behavior and that over one-quarter were missing a history of physical and social development. *Id.* at 33. In contrast, "the record of the I.Q. scores was clearly the most scrupulously kept record, and it appears to have been the most important one." *Id.* The almost exclusive use of I.Q. test results in EMR placement was contrary to a requirement in the California statute. See text accompanying notes 200 & 201 *supra*.

321. Slip op. at 38. Contrast this finding with that made in *Hobson*. See text accompanying notes 114 to 122 *supra*.

322. Slip op. at 38.

differences between males and females had been excised from early versions of the Stanford-Binet Intelligence Scale. Acknowledging that equalizing scores of minorities and whites might be difficult, it criticized testing experts for being "willing to tolerate or even encourage tests that portray minorities, especially blacks, as intellectually inferior."³²³

The court examined the possible explanations for the significantly disparate scores of blacks and whites on I.Q. tests. Rejecting genetic³²⁴ and socioeconomic³²⁵ explanations of the disparities in scores, the court turned to the hypothesis that cultural bias in the tests caused the disparities. It noted that versions of the Stanford-Binet and Wechsler scales used prior to the 1970's had been developed using only white children in the process of deriving norms against which all children would be measured. That these tests had been restandardized in the early 1970's to include a representative proportion of black children did not satisfy the court that they were valid for culturally different groups: "Mixing the populations without more does not eliminate any preexisting bias."³²⁶ The process failed to yield data that could be used to compare black and white children's performance on particular items.³²⁷ The court identified two other indicia of cultural bias. First, to the extent that black children were more likely to be exposed to nonstandard English, they would be handicapped in the verbal component of intelligence tests.³²⁸ Second, it posited that certain items were inherently biased against black children from culturally different environments when viewed from the perspective of the scoring

323. *Id.* at 41.

324. The defendants did not rely on this ground. Wilson Riles, superintendent of the California state department of education, stated that the genetic rationale was "reprehensible." *Id.* at 42. However, some state officials did testify that they would not rule out a genetic explanation, and one defense expert witness, psychologist Robert Gordon, believed that it was the most cogent reason for I.Q. score disparities between the races. *Id.*

325. The defendants contended that differences in scores resulted from the rearing of poor children, both black and white, in inadequate homes and neighborhoods. Although the court could accept that poverty might lead to academic deficiencies, it would not accept the theory that poverty resulted in mental retardation. Why, the court asked, would inadequate financial resources produce disproportionately more mildly retarded children — as those found in EMR classes — and not cause a disproportionate number of children with severe intellectual deficits? *Id.* at 44. It is possible that nutritional deficiencies in poor people affect intellectual functioning to a limited degree but are not serious enough to produce severe or profound retardation.

326. *Id.* at 46.

327. *Id.* at 122 n.64.

328. *Id.* at 47. However, the court admitted that black children also did significantly less well on some nonverbal tasks. *Id.* at 123 n.66.

criteria offered in the examiner's manual.³²⁹ The court concluded that "to the extent that a 'black culture' exists and translates the phenomenon of intelligence into skills and knowledge untested by the standardized intelligence tests, those tests cannot measure the capabilities of black children."³³⁰ The court charged that the tests were never designed to eliminate bias against black children and blamed test developers and users for assuming "in effect that black children were less 'intelligent' than whites."³³¹

The court had constructed an analytic web from which the defendants could not extricate themselves. By defining purposeful discrimination to mean the intent to segregate minority children into special, isolated classes, the court laid the groundwork for a decision in favor of the plaintiffs. It judged California's EMR program to be a substandard, stigmatizing means of education, a virtual prison from which black children could not easily escape.³³² It concluded that on the basis of intelligence tests mandated by the state EMR classes had been populated by a disproportionate number of minority children for a decade. The process by which the tests were chosen was haphazard, unthinking, and suspect, making "the inference of discriminatory intent . . . inescapable."³³³ The tests themselves were seen by both plaintiffs' and defendants' expert witnesses to be culturally biased, since none of them had been specifically validated for black children.³³⁴ Perhaps most damaging to the defendants'

329. The court chose one item for particular scrutiny. On one of the subtests of the Wechsler Intelligence Scale for Children, the examinee is asked, "What is the thing to do if a boy/girl much smaller than yourself starts to fight with you?" The scoring manual gives the most credit for behavior that avoids retaliation. According to one expert's testimony at trial, this question is answered incorrectly more than twice as often by young black children than white. *Id.* at 48. *But see* text accompanying notes 130 to 133 *supra*.

330. Slip op. at 50.

331. *Id.* at 45.

332. "Minority children . . . are isolated for virtually the entire school day from mainstream children, and E.M.R. classes provide an inferior education for children assigned there . . ." *Id.* at 84.

333. *Id.* at 88. The court stated:

The decision by the State Department of Education in 1969 to compel the use of standardized I.Q. tests . . . reveals the impermissible intent to discriminate. . . . [I]t had profound discriminatory effects. It doomed large numbers of black children to E.M.R. status, racially imbalanced classes, an inferior and "dead-end" education, and the stigma that inevitably comes from the use of the label "retarded." That impact was not only foreseeable but foreseen, and appropriate inferences must be drawn.

Id. at 85-86. *See also id.* at 86-87.

334. *Id.* at 88. The court was strident in its criticism of the state:

Defendants' complete failure to ascertain or attempt to ascertain the validity of the tests for minority children cannot be ignored. Rather

position was the court's critical perception of the state's conduct. Condemning the complacency and negligence of the state department of education in the face of explicit legislative concern about biased testing and disproportionate enrollment of minorities in EMR classes, the court charged that the state had not investigated the problems, had not inquired into significant variances in the racial and ethnic composition of classes for the retarded, and had failed to monitor implementation of protections imposed by the legislature to insure that placement decisions would be made on bases other than I.Q. tests.³³⁵ The court concluded that the state's conduct "must be seen as a desire to perpetuate the segregation of minorities in inferior, dead-end, and stigmatizing classes for the retarded."³³⁶

Thus the plaintiffs were held to have met their burden of proving discriminatory intent. The defendants could prevail only if explanations for their conduct passed muster under the most exacting of the equal protection tests. However, the court concluded, "Defendants can establish no compelling state interest in the use of the I.Q. tests nor in the maintenance of EMR classes with overwhelming disproportions of black enrollment."³³⁷

After finding for plaintiffs under both federal statutes and the Constitution,³³⁸ what remained for the court was to forge proper remedies. The court permanently enjoined the state from using any standardized intelligence tests to identify black children for EMR

defendants' actions resulting in the adoption of the I.Q. requirement and the short list of accepted I.Q. tests can only be explained as the product of the impermissible and scientifically dubious assumption that black children as a group are inherently less capable of academic achievement than white children.

Id.

335. *Id.* at 88-91.

336. *Id.* at 92.

337. *Id.* at 94. Defendants asserted that the tests prevented misclassification rather than caused it, that there were no real alternatives to I.Q. testing, and that testing helped promote fiscal planning. All these defenses were brushed aside as post hoc justifications. *Id.* What further damaged the defendants' case was that the state had earlier agreed to eliminate disparities for Hispanic children. *Id.* at 95; see text accompanying notes 166 to 173. The court found that the plaintiffs could also prevail under an intermediate standard of review applicable when a nonsuspect classification operates to the detriment of a disadvantaged class of persons. Such review would require that the school system's classification "substantially furthers the purpose" of providing education. Slip op. at 95-96. See *Berkelman v. San Francisco Unified School Dist.*, 501 F.2d 1264 (9th Cir. 1974).

338. The court also found that defendants had violated California's constitutional guarantee of equal protection but found no violation of its education code. Slip op. at 97, 131 n.110.

placement without first securing its approval.³³⁹ With regard to disproportionate placement, the state was ordered to monitor and eliminate overrepresentation by obtaining annual data documenting enrollment in EMR classes by race and ethnicity and by requiring each school district to prepare and adopt plans to correct significant imbalances.³⁴⁰ To remedy the harm to those children misidentified, the defendants were to reevaluate all black children then labeled as educably retarded without resort to any standardized intelligence tests that had not been approved by the court.³⁴¹ Finally, schools would have to draft individual education plans designed to return all incorrectly identified children to regular classrooms.³⁴²

Critique of Riles II

Riles II is an interesting mixture of judicial innovation and conservatism. The court conscientiously cited precedent at every step of its analysis, attempted to conform its decision to the existing law in its own circuit as well as to that of the Supreme Court, and carefully — if not always correctly — considered all the viewpoints presented in the overwhelming amount of testimony. In form and substance the decision is reminiscent of *Hobson v. Hansen*,³⁴³ decided a dozen years earlier. The concern for precedent was appropriate because Judge Peckham's decision in *Riles II* is just as revolutionary in 1979 as was Judge Wright's in 1967. While *Hobson*

339. In doing so it recognized the changes initiated by California since the issuance of the preliminary injunction and the complexity and risk of judicial interference in the administration of education. It also did not wish its condemnation of intelligence tests to be seen as the "final judgment on the scientific validity" of such devices. *Id.* at 102. But these concerns did not dissuade the court from holding the state responsible for its failure to properly assess and educate black children, and from fashioning remedies to halt both test abuse and disproportionate enrollment of blacks in EMR classes. The state board of education would have to petition the court, after determining that the tests they sought to use were not racially or culturally discriminatory, that they would be administered in a nondiscriminatory manner, and that they had been validated for the purpose of placing black children in EMR classes. The petition must be supported by statistical evidence submitted under oath and certification that public hearings had been held concerning the proposed tests. *Id.* at 104-05.

340. *Id.* at 105-06. The court required such plans for any school system in which the rate of black EMR enrollment was one standard deviation above the rate for white EMR enrollment. The court did not clarify what it meant by this requirement.

341. *Id.* at 106-07.

342. *Id.* at 107. Individual education plans are required for all handicapped children under Pub. L. No. 94-142. See 45 C.F.R. § 121a.340-.349 (1977). The court refused to order compensatory education for black disadvantaged children already in regular classes. Slip op. at 108.

343. 269 F. Supp. 401 (D.D.C. 1967). See text accompanying notes 79 to 110 *supra*.

proscribed tracking and the use of group achievement tests, *Riles II* for the first time permanently enjoined the use of the traditional and most respected tool of the professional psychologist's trade — the individually administered intelligence test. The case may also have signalled the end of separate classes for mildly retarded children whether populated by white or minority children. The court called EMR classes "educational anachronisms" that focused on labels rather than individual needs, isolated students, and relegated them to inferior education rather than providing remedial training designed to return them to regular classes containing nonhandicapped students.³⁴⁴

Of equal, though not such dramatic, importance was the court's attempt to further define the meaning of discriminatory intent under the equal protection clause in the post-*Davis* era. Perhaps even more significant, however, was the court's endeavor to give meaning to the ambiguous concepts of nondiscriminatory assessment and test validation. As a result, test publishers and users must now produce data showing that tests have validity for each discrete group for which they will be used. Standardization samples that include minority groups but do not yield separate norms for each race will not be judged valid. I.Q. test scores will have to be correlated with relevant measures, not simply scores derived from other standardized ability tests. One wishes that the court had gone further and guided examiners as to the level of validity — that is, the coefficient of correlation — required to make a test acceptable³⁴⁵ or how much disparity between minority and white correlations would be tolerated before a test was considered discriminatory. Nevertheless, the court performed a useful service by giving future litigants one judicial definition of the requirement that tests be "validated for the specific purpose for which they are used."³⁴⁶ From a procedural perspective, one may be surprised at the ease with which the court permitted the retroactive application of federal law to the facts in this case. Section 504 of the Rehabilitation Act was not passed until 1973, and its implementing regulations did not take effect until June 1977. Public Law 94-142 was not passed until 1975, and its regulations only took effect in October 1977. Yet the court permitted the plaintiffs to

344. Slip op. at 109. This was in sharp contrast to Judge Peckham's decision in *Riles I*. See text accompanying note 208 *supra*. No doubt this shift was stimulated by the least restrictive environment provisions of federal legislation protecting handicapped persons.

345. See text accompanying notes 254 to 256 *supra*.

346. 45 C.F.R. § 121a.532 (1977).

amend their complaint in a suit the operative facts of which originated in 1971.³⁴⁷

Of course, the most significant aspect of *Riles II* is the court's permanent injunction of the administration of individual intelligence tests to place black children in EMR classes. The ban was based almost entirely on the court's finding that the tests were culturally biased, and the persuasiveness of the court's opinion depends on the correctness of this finding. Regardless of whether one applauds or decries the result, however, there are infirmities in the court's analysis. *Riles II* defined an unbiased test as one that yields "the same pattern of scores when administered to different groups of people."³⁴⁸ Such a definition is psychometrically unsound. Tests are fair when they predict with equal accuracy, not with equal results, for all groups. The court's definition "eliminates a priori any possibility of real group differences on various psychological traits. . . ."³⁴⁹ The court rejected the possibility of genuine differences between black and white children on the basis of genetic inferiority and socioeconomic influences. Though the court rested its decision on the finding that the tests were culturally biased, it provided little hard data to support such a conclusion and was tentative in discussing it. In fact, the court's determination that the tests contain questions biased against poor black children is not uniformly accepted, and there is some data to suggest that whatever discrimination there is in tests, lower scores in blacks are not the result of content bias.³⁵⁰ The court was correct in criticizing test publishers for not adequately standardizing their instruments on discrete minority populations, but beyond that, its analysis of cultural discrimination is weak, and the issue is certainly not as settled as the court appears to think. Nevertheless, the court could only rest its holding on the data presented to it by the parties. The state's defense was made difficult by the lack of relevant studies on differential validity, the absence of systematic research concerning content bias, and its concession that cultural differences affected I.Q. scores.³⁵¹

The court's equal protection analysis deserves some attention. Its discussion of the intermediate standard of review — the most

347. *But see, e.g., Stemple v. Board of Educ.*, 464 F. Supp. 258 (D. Md. 1979), *appeal docketed*, No. 79-1208 (4th Cir. March 30, 1979).

348. Slip op. at 41.

349. Schmidt & Hunter, *supra* note 134, at 1. See notes 135 & 136 and accompanying text *supra*.

350. See text accompanying notes 127 to 133 *supra*.

351. Slip op. at 48.

recently developed form of judicial scrutiny under the fourteenth amendment³⁵² — relied almost entirely on one court of appeals opinion, *Berkelman v. San Francisco Unified School District*.³⁵³ There, the Ninth Circuit concluded that “where a nonsuspect classification . . . is alleged to operate to the detriment of a disadvantaged class or classes [e.g., black and Spanish-American students], neither ‘strict’ nor ‘minimal’ scrutiny provides useful guidance as a standard of review.”³⁵⁴ But, it is just this kind of interaction, present also in *Riles*, that could have led the court to employ the strict scrutiny test.

In *Riles* there was an explicit classification based on intelligence and an implicit classification based on race. While either alone probably does not raise the standard to strict scrutiny, the combination of the two might be so stigmatizing as to require defendants to show more than a rational relationship or even substantial congruence (as is required under the intermediate standard of review) between their practices and goals. De facto racial discrimination in the setting of de jure intelligence classification

compounds the psychological injury to the individual and introduces the factor of a discrete and insular group [T]he two levels of discrimination . . . inevitably reinforce each other, hastening the downward spiral of self-confidence and self-esteem. The potentially dangerous consequences generated by the creation of a racial group stereotype based on low intelligence gives greater force to the stigmatic effect.³⁵⁵

Thus, minority students who are labeled intellectually defective suffer a cumulative disadvantage. Society has historically subjected them to intentional and protracted discrimination on the basis of race; prevented them from marrying, voting, procreating, and attending school on the basis of their handicaps; and attenuated their constitutional rights relative to adults on the basis of their status as children.³⁵⁶ If, as one legal scholar has suggested, the equal

352. See note 184 *supra*.

353. 501 F.2d 1264 (9th Cir. 1974).

354. *Id.* at 1267.

355. *Intelligence Classifications*, *supra* note 8, at 672; cf. *Lora v. Board of Educ.*, 456 F. Supp. 1211, 1276 (“Where . . . the handicapped children are also from minority groups, the burden of showing a sufficient educational opportunity is especially high;” allegations of discrimination in the assessment and placement of emotionally disturbed black and hispanic students).

356. The most recent case reinforcing the notion that children are accorded less constitutional protection than adults is *Parham v. J.L. & J.R.*, 99 S. Ct. 2493 (1979).

protection clause, and the Constitution as a whole, were meant to be "participation-oriented [and] representation-reinforcing,"³⁵⁷ then there is a sound basis for arguing that mentally retarded minority children deserve the highest protection afforded by the Constitution.

Perhaps because the court believed that its revolutionary remedies precluded more adventuresome analysis, it lost the opportunity to rule that strict scrutiny of defendants' action was appropriate, not only because of the intent to discriminate but because their conduct impermissibly affected a suspect class of persons. This approach would not have been entirely novel.

Some commentators³⁵⁸ have suggested that intelligence, like race, national ancestry, and alienage, should be considered a suspect classification. Classifications already denominated as suspect are those that have affected persons burdened either with disabilities or a history of purposeful unequal treatment or who have been relegated to a position of political powerlessness — discrete and insular minorities who command extraordinary protection against the will of the majority.³⁵⁹ Intelligence classifications, it has been claimed, fall within that definition:

Like one's race, . . . a person's intelligence (at least with regard to innate capacity) is a matter over which he lacks control. . . . [A]n individual cannot wilfully raise or lower his IQ more than a few points on a scale. As it does with other minorities, moreover, the selective disadvantaging of less intelligent people tends to punish [those in] possession of the identifying characteristic by perpetuating a cycle of failure and defeat. . . . It has been found that those who are classified as "dull" often begin to exhibit the characteristics associated with such a classification, regardless of their actual mental capacity, and so are unnecessarily disadvantaged in life.³⁶⁰

While facially persuasive, the commentators' analogy of intelligence to race is faulty on several grounds. In some cases, an individual's I.Q. scores can vary widely across time,³⁶¹ and there is evidence that parents can train their children to perform better on I.Q. tests.³⁶² Further, the self-fulfilling prophecy theory implied in

357. Ely, *Toward a Representation-Reinforcing Mode of Judicial Review*, 37 MD. L. REV. 451, 471 (1978).

358. Shea, *supra* note 8, at 173; *Intelligence Classifications*, *supra* note 8, at 653-56.

359. See note 185 *supra*.

360. *Intelligence Classifications*, *supra* note 8, at 654-55.

361. Anastasi, *supra* note 7, at 329-32.

362. Whimbey, *You Can Learn to Raise Your IQ Score*, 9 PSYCH. TODAY 27 (1976).

the last sentence of the excerpt has been attacked as not supportable, at least with regard to students' performance in classrooms.³⁶³

In light of a comment by one Justice, it seems unlikely, however, that the Supreme Court will consider intelligence a suspect classification. In discussing whether sex was a suspect classification, Justice Brennan differentiated it from "such nonsuspect statuses as intelligence or physical disability."³⁶⁴ Moreover,

discrimination on the basis of intelligence is not usually directed at a discrete and insular minority. . . . [I]ntelligence varies along a spectrum. . . . [M]embers of the group will vary as arbitrary decisions about scales and cut-off points vary in particular cases. . . . Given the absence of discreteness, and . . . given the perceived lack of insularity, there is little concern that intelligence classifications deny effective political representation.³⁶⁵

If one views low intelligence as a handicap, however, it may be easier to argue that intelligence is a suspect classification.³⁶⁶ Using the *Rodriguez* definition, the contention is that handicapped persons are by definition "saddled with disabilities,"³⁶⁷ are politically powerless,³⁶⁸ have suffered a history of purposeful discrimination,³⁶⁹ and are discrete and insular minorities possessing randomly

363. The theory gained national recognition as the result of a study in which it was allegedly demonstrated that teachers who were erroneously informed that some of their students were "late bloomers" began treating these normal students as bright. R. ROSENTHAL & L. JACOBSON, *PYGMALION IN THE CLASSROOM* (1970). The study has been severely criticized for its methodological deficiencies. Thorndike, Book Review, 5 *AM. EDUC. RESEARCH J.* 708 (1968).

364. *Frontiero v. Richardson*, 411 U.S. 677, 686 (1973) (plurality opinion).

365. *Intelligence Classifications*, *supra* note 8, at 655.

366. See, e.g., Burgdorf & Burgdorf, *A History of Unequal Treatment: The Qualifications of Handicapped Persons as a "Suspect Class" Under the Equal Protection Clause*, 15 *SANTA CLARA L. REV.* 855 (1975); Burt, *Judicial Action to Aid the Retarded*, in 2 *ISSUES IN THE CLASSIFICATION OF CHILDREN* 293, 303-12 (N. Hobbs ed. 1975); Shaman, *Persons Who Are Mentally Retarded: Their Right to Marry and Have Children*, 12 *FAM. L.Q.* 61, 66-70 (1978); *Segregation*, *supra* note 8, at 1236. See also Note, *Mental Illness: A Suspect Classification?*, 83 *YALE L.J.* 1237 (1974).

367. *San Antonio Ind. School Dist. v. Rodriguez*, 411 U.S. 1, 28 (1973).

368. This is especially true of retarded persons and those considered to be mentally ill who have been "denied the right to vote by express provisions in state constitutions and statutes." Burgdorf & Burgdorf, *supra* note 366, at 906.

369. The strongest evidence of this is compulsory sterilization laws and denial of access to public schools. *Fialkowski v. Shapp*, 405 F. Supp. 946, 959 n.9 (E.D. Pa. 1975); *Pennsylvania Ass'n for Retarded Children v. Pennsylvania*, 343 F. Supp. 279, 293-95 (E.D. Pa. 1972).

occurring immutable characteristics and as such are readily identifiable:

A group of persons labeled as *retarded* have in this society been subjected to discrimination as brutal and dehumanizing as those imposed on the slave population. . . . This society's willingness to confine retarded human beings in . . . terrible warehouses, and to shut these warehouses from sight by their remote rural locations, is eloquent testimony that "retarded" persons — like slaves . . . — have been made nonpersons. If further confirmation is needed to demonstrate the brutal dehumanizations this society has practiced on retarded persons as a class, it is enough to consider the compulsory sterilization laws enacted during the first decades of this century with the false and unjustified promise to cleanse this country of "generations of imbeciles."

Retarded children . . . are as a rule the victims of birth or developmental defects for which they can bear no just blame. And, as a class, they are subjected to galling and brutal social discriminations.³⁷⁰

The contention that handicapped persons should be entitled to extraordinary protection under a suspect class status has received some favorable judicial consideration. The North Dakota Supreme Court, hearing a claim by a young quadruple amputee with other physical and sensory defects that she had a right to an education in the public schools, called the plaintiff's multiple handicaps "just the sort of 'immutable characteristic determined solely by accident of birth' to which the 'inherently suspect' classification would be applied."³⁷¹ A federal court in New York acknowledged that "emotionally disturbed children might be characterized as a suspect class. . . ."³⁷² When two severely retarded, multiply handicapped

370. Burt, *supra* note 366, at 303-05. "Handicapped persons are a distinct minority, frequently isolated from the rest of society. They bear the brunt of social prejudice and tend to be . . . cut off from the political process. . . . [M]any handicaps are immutable characteristics occurring on a random basis, punishment for which is 'illogical and unjust.'" Burgdorf & Burgdorf, *supra* note 366, at 907.

371. *In re G.H.*, 218 N.W.2d 441, 447 (N.D. 1974) (citing *Frontiero v. Richardson*, 411 U.S. 677 (1973) (plurality opinion)). Much of that quotation, however, comes from *Weber v. Aetna Cas. & Sur. Co.*, 406 U.S. 164, 175, in which the Supreme Court considered but did not decide whether illegitimacy was a suspect class. Reliance on *Frontiero* was probably misplaced given Justice Brennan's remark that physical disability was not a suspect class. *Frontiero v. Richardson*, 411 U.S. 677, 686 (1973) (plurality opinion).

372. *Lora v. Board of Educ.*, 456 F. Supp. 1211, 1275 (E.D.N.Y. 1978).

plaintiffs,³⁷³ seeking a ruling that they had been deprived of an appropriate education, asserted that their claims should be strictly scrutinized, a federal district court found "a certain immediate appeal to [their] argument."³⁷⁴ Although it never decided the issue, the court agreed that "retarded children are precluded from the political process and have been neglected by the state legislatures. Moreover the label 'retarded' might bear as great a stigma as any racial slur."³⁷⁵

However, one can empathize with Judge Peckham's neglect of this line of analysis. Even though there is "ample evidence to show that retarded [and handicapped] persons are an unjustly stigmatized, politically vulnerable group and that state policies . . . have fanned popular social opprobrium,"³⁷⁶ those advocating that handicapped persons' equal protection claims should be strictly scrutinized face increased resistance from the Supreme Court to recognizing new suspect classifications. In the last six years the Court has denied suspect classification status to wealth,³⁷⁷ sex,³⁷⁸ and illegitimacy,³⁷⁹ all of which had appeared prime candidates for that designation in earlier cases.

Thus, while the case for handicapped persons is strong, they may not have suffered the kind of profound and encompassing injuries that require "extraordinary protection from the majoritarian political process."³⁸⁰ Further, the handicapped may be too amorphous a group to meet the test of definability that is required for consideration as a suspect class.³⁸¹ Perhaps the more discrete handicap of retardation would meet that criterion, but even that category may be too large, for it comprises those who range from

373. One was 21 years old with a reported mental age of 19 months, the other 12 years old with a reported mental age of 15 months. *Fialkowski v. Shapp*, 405 F. Supp. 946, 948 (E.D. Pa. 1975).

374. *Id.* at 958.

375. *Id.* at 959. *But see* New York Ass'n for Retarded Children v. Rockefeller, 357 F. Supp. 752, 762-63 (E.D.N.Y. 1973).

376. Burt, *supra* note 366, at 305.

377. *San Antonio Ind. School Dist. v. Rodriguez*, 411 U.S. 1 (1973).

378. *See* General Elec. Co. v. Gilbert, 429 U.S. 125 (1976); *Geduldig v. Aiello*, 417 U.S. 484 (1974).

379. *Mathews v. Lucas*, 427 U.S. 495 (1976). Though the Court agreed that illegitimacy, like race, was a characteristic one did not freely choose nor did it bear any relation to the individual's ability to participate in and contribute to society, it found that, unlike race, illegitimacy was not obvious and overtly recognizable, nor was the history of legal and political discrimination against illegitimates as severe and pervasive as that against racial minorities. *Id.* at 505-06.

380. *San Antonio Ind. School Dist. v. Rodriguez*, 411 U.S. 1, 28 (1973).

381. *See id.* at 22.

barely functioning persons with unmeasurable intelligence, *i.e.*, less than an I.Q. score of 10, to those whose I.Q. score is two standard deviations below the mean, *i.e.*, about an I.Q. score of 70. Possibly, the suspect classification status could apply to the profoundly retarded,³⁸² but that would be of no help to plaintiffs in such cases as *Riles*, where the challenge is to the appropriate placement of educably retarded children who have suffered less at the hands of the state than the severely impaired.

Finally, what is perhaps most instructive about *Riles II* is the court's harsh condemnation of the California state department of education. The department's unsystematic, hasty, and unsound methods in developing the list of approved I.Q. tests, its failure to monitor adequately the placement of minorities in EMR classes, and its disregard of legislative concern about minority overrepresentation in inferior programs drew angry and extensive criticism. *Riles II* may be another demonstration that conduct on the part of government officials, perceived as consciously disregarding the rights of citizens, may lead to landmark decisions that have more profound consequences than simply the correction of the irregularities that provoked them.

LITIGATION GENERATED BY STATUTES

It was earlier noted that an intense period of litigation may culminate in the passage of legislation that incorporates the principles developed piecemeal in case law and applies them to broad classes of persons. But, as exemplified in *Riles II*, legislation itself may serve as a springboard to future litigation, as parties seek to define, implement, and enforce its provisions. Federal and state statutes may evoke what may be called "second generation" issues. With access to a free public education an established principle, the questions are how broadly the right may be invoked and what particular rights are embodied in the general assurance.

In *Frederick L. v. Thomas*,³⁸³ for example, a federal district court interpreted the right to education to require school systems to engage in massive screening and subsequent individual psychological

382. As with the plaintiffs in *Fialkowski v. Shapp*, 405 F. Supp. 946 (E.D. Pa. 1975), see notes 373 & 374 *supra*, "[A] court might hold 'intelligence' to be a suspect class but might invalidate its use only when it imposes the most gross deprivations on the most vulnerable members of the class." Burt, *supra* note 366, at 311-12.

383. 419 F. Supp. 960 (E.D. Pa. 1976), *aff'd*, 557 F.2d 373 (3d Cir. 1977), *remanded for further proceedings*, 578 F.2d 513 (3d Cir. 1978) (granting of defendants' motion for abstention affirmed so that state court could consider claim for damages).

evaluations for learning disabled children. *Frederick L.* was a class action against the Philadelphia school system charging that it had failed to provide learning disabled children with appropriate educational programs as required by Pennsylvania's special education laws and the federal Constitution.³⁸⁴ Expert witnesses for both sides testified that the incidence of learning disability in the defendant school system was three percent, and on that basis estimated that of the 263,000 students enrolled in the Philadelphia schools, approximately 7,900 were learning disabled.³⁸⁵ Only 1300 were being served in special programs.³⁸⁶ The defendants contended that the remaining 6,600 were properly "mainstreamed" in regular classes and did not need special services, as they were referred neither by their parents nor their teachers for additional help. The plaintiffs responded that the school could not presume that such children were being properly served unless the determination was made on the basis of an appropriate evaluation. They claimed that failure to perform comprehensive assessments resulted in the functional exclusion of the 6,600 children, and thus denied them equal educational opportunity under the fourteenth amendment's equal protection clause.

The court found for the plaintiffs on the basis of Pennsylvania's special education laws and regulations. Although it agreed with the defendants that learning disabled children could be offered minimal special services and could be provided nothing more than the monitoring of academic performance in regular classrooms, the court refused to find that the children could be placed in regular programs without identification. "There must first be a diagnosis of the child's exceptionality, and then a determination that mainstreaming under prescribed conditions will be an appropriate placement."³⁸⁷ The court ordered the school district to develop a procedure that was "reasonably calculated to identify all of its learning disabled students within the near future."³⁸⁸

The court noted that parent referrals may not be the most effective means of identifying learning disabled children because some parents may not know enough about handicaps to realize that their children need special instruction. It also was concerned that

384. *Id.* at 962.

385. *Id.* at 963.

386. *Id.* at 974.

387. *Id.* at 973.

388. *Id.* at 974.

teachers could not always serve as adequate referral sources. Testimony had shown that teachers often neglected to request special services for nondisruptive but academically deficient students and that even disruptive pupils were not referred if teachers believed that there was no existing suitable program for them.³⁸⁹ The court left to the school district the formulation of a plan for the identification of all learning disabled students. It would seem, however, that only system-wide screening of the entire student body followed by individual testing of those subsequently found to be possibly learning disabled would suffice.³⁹⁰ The implications of *Frederick L.* are significant. It places further pressure on already embattled and financially exhausted school systems; it almost inevitably restricts school psychologists to the role of tester; and, ironically, while other cases have curbed extensive reliance on tests, it expands their use greatly.

In light of *Frederick L.* it is interesting to note that Congress, when it was considering Public Law 94-142, was concerned that the vague definition of "specific learning disability" would lead to the overinclusion of many children, thus inflating the numbers of handicapped children for whom federal reimbursement was available.³⁹¹ To prevent overinclusion it placed an upper limit of two percent on the number of all children that could be labeled as learning disabled, to be in effect until the Commissioner of Education published regulations redefining the condition, established specific criteria for determining its existence, and developed procedures for arriving at the diagnosis. The proposed regulations provided that (1) it was not presently possible to specify all the components of each learning disability; (2) no generally accepted diagnostic instruments could be used appropriately with all learning disabled children; (3) it was not possible to list specific standardized

389. *Id.* at 973.

390. *Id.* at 965.

391. For purposes of the Act, Congress defined learning disability in the traditional words of the exclusionary description found in standard texts:

The term "children with specific learning disabilities" means those children who have a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which disorder may manifest itself in imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. Such disorders include such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Such term does not include children who have learning problems which are primarily the result of visual, hearing, or motor handicaps, of mental retardation, or of environmental, cultural, or economic disadvantage.

20 U.S.C. § 1411 (1978).

instruments to be used in the evaluation of such children; and (4) regulation of the diagnostic process would be extremely extensive and impractical.³⁹²

The final regulations retain the original definition with only one variation,³⁹³ and require that any determination of learning disability be grounded in a severe discrepancy between achievement and intellectual ability in one of seven areas.³⁹⁴ Further, they require that, in addition to the assessment of I.Q. and educational performance, at least one member of the evaluation team observe the child's academic behavior in the regular classroom setting.³⁹⁵ The regulations do little to advance the theory, assessment, or diagnosis of learning disabilities. Diagnosis is still clearly tied to measured intelligence as the criterion for ability, with all the problems that that entails.³⁹⁶

Another case brought on by legislation is *Lora v. Board of Education*,³⁹⁷ in which black and Hispanic emotionally disturbed students claimed that inadequate and improper psychoeducational evaluations led to racially discriminatory referral and assignments to special day schools that were removed from the mainstream of regular educational settings. The assessment of children considered to be candidates for placement in these schools was the responsibility of evaluation teams consisting of psychologists, learning disability specialists, social workers, psychiatrists, and guidance counselors.³⁹⁸ Plaintiffs alleged that notwithstanding this procedure, the defendants placed excessive reliance on tests that were culturally biased.

392. 41 Fed. Reg. 52,404-05 (1976).

393. 45 C.F.R. § 121a.5(b)(9) (1977). It excluded academic deficiency based on emotional disturbance as a learning disability.

394. *Id.* § 121a.541. The areas are oral expression, listening comprehension, written expression, basic reading skill, reading comprehension, mathematics calculation, and mathematics reasoning.

395. *Id.* § 121a.542.

396. *E.g.*, J. SALVIA & J. YSELDYKE, *supra* note 4; Coles, *The Learning Disabilities Test Battery: Empirical and Social Issues*, 48 HARV. EDUC. REV. 313 (1978).

397. 456 F. Supp. 1211 (E.D.N.Y. 1978).

398. The evaluation teams, sensitive to the possibilities of cultural and racial bias, used what the defendants called a "pluralistic" model of assessment. Under this system, a social worker obtained background information concerning the educational and developmental history of the child and his family environment, an educational evaluator assessed the child's academic strengths and weaknesses, a psychiatrist interviewed the child, where appropriate, a neurologist evaluated the child's neurological functioning, and a psychologist administered a series of tests to assess intellectual and emotional functioning. In addition, some children were observed in a simulated learning environment (the diagnostic classroom). To prevent individual biases from predominating, the evaluation team met to decide jointly on classification

Invoking the due process and equal protection clauses,³⁹⁹ Title VI of the Civil Rights Act of 1964,⁴⁰⁰ Public Law 94-142,⁴⁰¹ and section 504 of the 1973 Rehabilitation Act,⁴⁰² the court analyzed the plaintiffs' claims using a right to treatment theory.⁴⁰³ It stated that "since proper evaluation is . . . central to acceptable special education, a program falling substantially below minimum established standards would constitute a violation of the right to treatment."⁴⁰⁴ The court found that some of the defendants' assessment practices fell short of those standards; however, the court did not condemn the tests used or even examine claims of test bias. Rather, it considered the possibility of discrimination only insofar as assessments were performed by nonminority professionals. It focused its attack primarily on the long waiting lists of children to be evaluated and placed, the lack of systematic annual review of students, and the transfer of students from one special education program to another without a full diagnostic evaluation. It found that the school violated requirements in Public Law 94-142 and section 504 for triennial reevaluations to determine whether children should be retained or integrated in the mainstream of regular education.⁴⁰⁵ Thus, as in the early southern desegregation cases, the court did not closely scrutinize the soundness of the instruments used to assess minority children for placement. Instead, it condemned omissions in established evaluation procedures and the lack of personnel to carry them out. Like *Frederick L.*, the ultimate result was to mandate, rather than restrict, further testing.

Frederick L. and *Lora* are only two of several cases that have been brought to the courts since the passage of Public Law 94-142, the Rehabilitation Act, and complementary state special education laws.⁴⁰⁶ Not all of this litigation concerned psychological testing, but there is little doubt that assessment practices will once again be the subject of significant litigation.

and placement. Finally, a separate committee reviewed the evaluation unit's decision in a meeting to which parents were invited. *Id.* at 1237.

399. *Id.* at 1232 (citing *Hobson* and *Riles I* as support).

400. 42 U.S.C. § 2000d (1978).

401. 20 U.S.C. §§ 1401-1461 (1978).

402. 29 U.S.C. §§ 701-794 (1975).

403. 456 F. Supp. at 1214-79.

404. *Id.* at 1285.

405. *Id.* at 1285-87, 1291.

406. See, e.g., *Boxall v. Sequoia Union High School Dist.*, 464 F. Supp. 1104 (N.D. Cal. 1979); *Stemple v. Board of Educ.*, 464 F. Supp. 258 (D. Md. 1979); *Howard S. v. Friendswood Ind. School Dist.*, 454 F. Supp. 634 (S.D. Tex. 1978); *Stuart v. Nappi*, 443 F. Supp. 1235 (D. Conn. 1978); *Kruse v. Campbell*, 431 F. Supp. 180 (E.D. Va. 1977), *vacated and remanded*, 434 U.S. 808 (1977).

PERSONALITY ASSESSMENT

Though personality tests may invade the fundamental right of privacy, in only one case,⁴⁰⁷ *Merriken v. Cressman*,⁴⁰⁸ has the effect of that right on the use of such tests been questioned. The case had its origins in 1970 when a survey ordered by the Commissioners of Montgomery County, Pennsylvania and conducted by a company called Scientific Resources revealed that many children in the county were using drugs. Most of these children, the study claimed, possessed certain common characteristics. For example, eighty percent of the identified drug abusers allegedly felt estranged from their families. On the basis of such data Scientific Resources suggested that the County Drug Commission sponsor a drug prevention program called Critical Period of Intervention (CPI). All three of the county school districts agreed to participate in the program.

There were to be two phases in the study — identification and therapy-intervention. In the first phase, questionnaires were to be

407. Why there is only a single case in this area is not clear. Although ability testing is more pervasive than personality testing, school systems do use figure drawings, projective tests, and paper and pencil personality inventories, especially in the assessment of emotional and behavioral difficulties. *E.g.*, *Lora v. Board of Educ.*, 456 F. Supp. 1211, 1237 (E.D.N.Y. 1978). Figure drawing tests include the House-Tree-Person Test and the Draw-A-Person. Projective tests include the Thematic Apperception Test (in which persons are asked to tell stories about supposedly ambiguous black and white drawings), the Rorschach Inkblot Test, and the Sentence Completion Test. Paper and pencil measures include the California Personality Inventory and the Edwards Personal Preference Schedule. *See generally* A. ANASTASI, *supra* note 7, at 493-616; J. SALVIA & J. YSELDYKE, *supra* note 4, at 373-80. Perhaps because they do not impinge on racial and cultural interests they are less vulnerable to challenge. In any event, as concern for privacy in public institutions heightens and the issue of discriminatory evaluation becomes defused through legislative regulation, there is a greater likelihood of increased scrutiny of the administration of tests that are designed to pierce surface behavior and uncover deep-seated emotional conflicts. This may be especially true when the tests themselves inquire as to the more private aspects of parent-child relationships. The Supreme Court has long recognized the primacy of parent-child relationships as against state intervention. *E.g.*, *Parham v. J.L. & J.R.*, 99 S. Ct. 2493 (1979); *Planned Parenthood of Cent. Mo. v. Danforth*, 428 U.S. 52 (1976); *Wisconsin v. Yoder*, 406 U.S. 205 (1972); *Prince v. Massachusetts*, 321 U.S. 158 (1944); *Pierce v. Society of Sisters*, 268 U.S. 510 (1925); *Meyer v. Nebraska*, 262 U.S. 390 (1923). In that light, the *Merriken* case can serve as a painful lesson to mental health professionals who work in school settings. The case has stimulated ample discussion in the legal and psychological literature. Bersoff, *Professional Ethics and Responsibilities: On the Horns of a Dilemma*, 13 J. SCH. PSYCH. 359 (1975); Moskowitz, *Parental Rights and State Education*, 50 WASH. L. REV. 623 (1975); 2 FORDHAM URB. L.J. 599 (1974); 13 J. FAM. L. 636 (1973-74); 27 VAND. L. REV. 372 (1974). Much of the discussion in this section is found in somewhat altered form in Bersoff, *supra* note 247, at 363.

408. 364 F. Supp. 913 (E.D. Pa. 1973).

given to eighth grade students and their teachers so that certain students, deemed potential drug abusers, could become part of the therapeutic program.⁴⁰⁹ The second phase of the study was intervention. When the CPI staff had analyzed all the results, they would compile a list of children who had significant potential for becoming drug abusers. The list would then be given to the school superintendent, who would organize a joint effort among guidance counselors, teachers, school psychologists, and others to provide group therapy programs. One of these was called the Guided Group Interaction, a mandatory program to which the identified students would be assigned. One of its purposes was to use the peer group as "a leveler or equalizer insuring that its members do not stray too far from its ranks."⁴¹⁰ When the program was first developed, the school system did not intend to obtain parental consent for their children's participation.⁴¹¹ The study contained no provision for student consent.⁴¹²

409. Teachers were asked to identify pupils who most and least fit eight descriptions of anti-social behavior. For example, one description was "This pupil makes unusual or inappropriate responses during normal school activity." The student form was to be somewhat lengthier. First, students would be asked to assess their own behavior, for example, to state which of the following statements best described themselves: (1) Someone who will probably be a success in life; (2) One who gets upset when faced with a difficult school problem; (3) Someone who has lots of self-confidence; (4) A student who has more problems than other students. In the next part of the questionnaire they would be asked questions about their relationships with their parents and the behavior of their parents, to indicate, for example, whether one or both parents "tell me how much they love me" or "make me feel unloved" or "seem to regret that I am growing up and spending more time away from home." Finally, the students would select from among their classmates those who fit certain descriptive statements similar in kind to the ones given the teachers. *Id.* at 913, 916, 918.

410. *Id.* at 916.

411. The researchers, however, did plan to send the following letter to each parent:
Dear Parent:

This letter is to inform you that, this fall, we are initiating a Drug Program called "Critical Period of Intervention" (CPI). The aim of this program is to identify children who may be susceptible to drug abuse and to intervene with concrete measures to help these children. Diagnostic testing will be part of this program and will provide data enabling the prevention program to be specific and positive.

... We ask your support and cooperation in this program and assure you of the confidentiality of these studies. If you wish to examine or receive further information regarding the program, please feel free to contact the principal in your school. If you do not wish to participate in this program, please notify your principal of this decision. We will assume your cooperation unless otherwise notified by you. . . .

Id. at 917.

412. *Id.* at 914.

Sylvia Merriken, the mother of one of the intended participants in the study, and a therapist in a drug and alcohol rehabilitation center, complained to the principal of her son's school and to the school board. The American Civil Liberties Union agreed to represent Mrs. Merriken to seek an injunction in federal district court to prevent the school system from carrying out its plans. The ACLU argued on her behalf that the program would violate the constitutional rights of both Mrs. Merriken and her son. The court granted a temporary injunction prohibiting the county from implementing its proposal until the litigation was completed. Two of the three schools in Montgomery County discontinued their participation, leaving only the Norristown system, which Mrs. Merriken's son attended, as a defendant.

When the suit began, the school system offered to change its procedure so that written parental consent would be required. As another attempt to compromise, the school modified the test so that students who did not want to be included could return an uncompleted protocol. But the proposal contained no provision for student consent, and no data were to be provided from which students could make an informed choice about participating.⁴¹³

Of the many constitutional challenges Mrs. Merriken made the court entertained seriously only the one based on privacy.⁴¹⁴ It found that the highly personal nature of the questionnaire disrupted family associations and interfered with the right of the mother to rear her child. It stated that there "is probably no more private a relationship, excepting marriage, which the Constitution safeguards than that between parent and child. This Court can look upon any invasion of that relationship as a direct violation of one's Constitutional right to privacy."⁴¹⁵ Although there was no Supreme Court precedent, the district court declared that privacy was entitled to as much constitutional protection as free speech. But who possessed this right — the student, the parents, or both? Although the court seemed ready to answer that question when it declared that "the fact that students are juveniles does not in any way invalidate their right to assert their Constitutional right to privacy,"⁴¹⁶ it avoided reaching the question whether the lack of consent by

413. *Id.* at 914-15.

414. The plaintiffs claimed that the CPI Program would violate freedom of religion, speech, assembly, and the privilege against self-incrimination, as well as the right to privacy. *Id.* at 917.

415. *Id.* at 918.

416. *Id.*

children to the invasion of their privacy would be sufficient to invalidate the research:

In the case at Bar, the children are never given the opportunity to consent to invasion of their privacy, only the opportunity to refuse to consent by returning a blank questionnaire. Whether this procedure is Constitutional is questionable, but the Court does not have to face that issue because the facts presented show that the parents could not have been properly informed about the CPI Program and as a result could not have given informed consent for their children to take the CPI Test.⁴¹⁷

The court evaded two important issues: whether the failure to secure the child's consent was independently sufficient to discredit constitutionally the program and whether parents as guardians could waive their children's constitutional rights by consenting for them. Instead, the court concentrated on Mrs. Merriken's own right of privacy and found that she was unable to give genuinely informed consent to the invasion of her personal life because the parental permission letter was so inadequate. The court deridingly compared the letter to a Book-of-the-Month Club solicitation in which a recipient's silence could be construed as agreement.⁴¹⁸ The letter was also criticized because it was a selling device that convinced parents to allow children to participate, not, as it properly should have been, an objective document telling parents of the potentially negative features and dangerous aspects of the program. The court then admonished school officials:

The parents are not aware of the consequences [of participating] and there is no substitute for candor and honesty in fact, particularly by the school board who, as the ultimate decision maker as far as the education of our children is concerned, should give our citizenry a more forthright approach. The attempt to make the letter requesting consent similar to a promotional inducement to buy, lacks the necessary substance to give a parent the opportunity to give knowing, intelligent, and aware consent.⁴¹⁹

Persons may, of course, waive their constitutional rights, but it is fundamental that such waivers must be voluntary, knowing,

417. *Id.* at 919.

418. *Id.* at 914.

419. *Id.* at 920.

intelligent, and made with awareness of possible adverse consequences.⁴²⁰ Mrs. Merriken had the right to waive her right of privacy by consenting to the research and intervention program, but because the request was little more than huckstering, it lacked the necessary substance to afford her the opportunity to consent meaningfully to the examination of her personal life.

There were other problems with the program. The promotional letter promised confidentiality, but the program contemplated the development of a "massive data bank"⁴²¹ and the dissemination of data relating to specific, identifiable students to school superintendents, principals, guidance counselors, coaches, social workers, PTA members, and school board members. And, even if the school system had been more circumspect and had constructed means by which the data were less widely distributed (or not distributed at all), no promise of confidentiality could take precedence over a subpoena compelling the disclosure of the material to law enforcement officials. As the court warned,

[T]here is no assurance that should an enterprising district attorney convene a special grand jury to investigate the drug problem in Montgomery County, the records of the CPI Program would remain inviolate from subpoenas and that he could not determine the identity of children who have been labeled by the CPI Program as potential drug abusers.⁴²²

Parents were not at all informed of this possibility.

Compounding the other problems was the fact that the identification instruments did not possess enough psychometric soundness to overcome the hazards that may have flowed from their use. While there could have been considerable harm done to children correctly identified, the court was particularly concerned about those children incorrectly identified. In a statement that should raise the anxiety level of psychologists, it said, "When a program talks about labeling someone as a particular type and such a label could remain with him for the remainder of his life, the margin of error must be almost nil."⁴²³ The court acknowledged that if the program had

420. *E.g.*, *Fuentes v. Shevin*, 407 U.S. 67 (1971); *Brady v. United States*, 397 U.S. 742 (1970).

421. 364 F. Supp. at 916.

422. *Id.*

423. *Id.* at 920.

demonstrated adequate public need and had restricted itself to a minimal invasion of privacy, it might have passed constitutional scrutiny. But, after studying all of the evidence, the court struck the balance in favor of the individual:

[T]he reasons for this are that the test itself and the surrounding results of that test are not sufficiently presented to both the child and the parents, as well as the Court, as to its authenticity and credibility in fighting the drug problem in this country. There is too much of a chance that the wrong people for the wrong reasons will be singled out and counselled in the wrong manner.⁴²⁴

ACCOUNTABILITY THROUGH ACCESS TO RECORDS

In the past test protocols — questions and answer sheets — were guarded by publishers and administrators to prevent public disclosure. Members of the American Psychological Association (APA) could have been expelled from the organization or suffered a lesser sanction if they violated the APA Code of Ethics provision that limits access to psychological tests and other assessment devices to “persons with professional interests who will safeguard their use.”⁴²⁵ Recently, however, test materials have become more accessible to examinees, in part because of the Family Education Rights and Privacy Act (FERPA), perhaps better known as the “Buckley Amendment.”⁴²⁶

FERPA provides in part that educational institutions receiving federal funds under any program administered by the United States Office of Education must allow parents access to records directly related to their children and an opportunity to challenge those records at a hearing.⁴²⁷ Portions of FERPA⁴²⁸ were incorporated in Public Law 94-142 to provide parents of handicapped children with the right to inspect and review all education records pertaining to

424. *Id.* at 921.

425. Am. Psychol. Ass'n, *Ethical Standards of Psychologists*, 18 AM. PSYCH. 56, 59 (1963).

426. In 1974 Senator James Buckley offered an amendment to Part C of the General Education Provisions Act (title V of Public Law 93-380), which passed with little debate. Its consequences were not fully realized until later. Pub. L. No. 93-380, § 513, 88 Stat. 571 (codified at 20 U.S.C. § 1232g (1978)), *amended by* Family Education Rights and Privacy Act, Pub. L. No. 93-568, § 2(a), 88 Stat. 1858 (1978).

427. 20 U.S.C. § 1232g(a)(1)(A) (1978).

428. The basic provisions of FERPA are more fully developed in the implementing regulations. 45 C.F.R. § 99.1-.67 (1978).

the identification, evaluation, and placement of their children in special classes.⁴²⁹

Whether test protocols themselves are accessible under these laws is not yet clear. Education records are defined by FERPA regulations as documents directly related to a student and maintained by an educational agency or institution or by a party acting for it.⁴³⁰ Thus, while they may not necessarily bear on education, the records may include, for example, psychiatric and psychological reports received from other agencies. They certainly would include documents created by education agency employees such as the school psychologist. However, excluded from the definition of a record are papers "in the sole possession of the maker thereof, and . . . not accessible or revealed to any other individual . . ."⁴³¹

Although protocols themselves are usually not shown by the test administrator to anyone else, information from them usually is. Most often the only tangible shared evidence of a psychological assessment is the report that follows testing. Psychologists usually attend a case conference of school personnel to discuss diagnoses and recommendations for placement, at which they disclose the results of testing and give examples of responses.⁴³² Even if psychologists maintain the individual test records and responses in their offices, if the information from them is communicated orally to others attending the case conferences, it would seem that protocols can be considered records because they have been "revealed" to others. As such, they are accessible under FERPA. Such an interpretation is supported by the statement issued by Senators Buckley and Pell, joint sponsors of FERPA: "[I]f a child has been labeled mentally or otherwise retarded and put aside in a special class or school, parents would be able to review materials in the record which led to this institutional decision . . . to see whether these materials contain inaccurate or erroneous evaluations about their child."⁴³³ Section 121a.562 of the regulations implementing Public Law 94-142 requires that schools permit parents to inspect and review their children's records collected, maintained, and *used* by the education agency in its special education decisionmaking.⁴³⁴

429. 20 U.S.C. § 1415(b)(1)(A) (1978); 45 C.F.R. § 121a.502 (1977); *id.* § 121a.562-.569.

430. 45 C.F.R. § 99.3 (1978).

431. *Id.* § 99.3(b)(1).

432. *See, e.g.,* *Lora v. Board of Educ.*, 456 F. Supp. 1211, 1238 (E.D.N.Y. 1978).

433. Joint Statement in Explanation of Buckley/Pell Amendment, 120 CONG. REC. 39862 (1974), *reprinted in* 40 Fed. Reg. 1213 (1975).

434. 45 C.F.R. § 121a.562(a) (1977).

Psychologists' tests and test results are almost always used in this process and thus should be accessible.

The accountability inherent in the complaint procedures under Public Law 94-142 makes disclosure of test questions possible in another way. It can be anticipated that attorneys representing parents (or the school system if parents introduce the results of an independent evaluation)⁴³⁵ will not simply accept at face value the interpretations of test results presented in psychologists' reports. Under cross-examination psychologists very likely will be asked for the factual bases underlying their conclusions. For example, if a psychologist testifies that part of the evidence substantiating a diagnosis of mental retardation is a scale score of five on the Information subtest of the WISC-R,⁴³⁶ a knowledgeable cross-examiner will inquire as to the student's separate responses in an attempt to cast doubt that those responses were the kind given by genuinely retarded children, but were more like those of a culturally different child with normal intelligence. In such an examination the questions would have to be revealed. Psychologists refusing to comply with the request could severely damage the position of the party for whom they are testifying, since the hearing examiner would be denied access to information useful in arriving at a decision. Perhaps recognizing the potential harm to clients if psychologists refuse to testify, the APA revised its Code of Ethics and now simply cautions against "imparting unnecessary information which would compromise test security" but permits psychologists to "provide requested information that explains the basis for decisions that may adversely affect person[s] examined."⁴³⁷

A case recently decided by the Supreme Court, *Detroit Edison Co. v. NLRB*,⁴³⁸ is relevant to the question whether examinees may have access to test stimuli and their responses but does not offer any clear-cut answer. In that case a union alleged that employees at Detroit Edison were wrongfully denied promotion after taking

435. Under Public Law 94-142 parents have the right to obtain an independent psychoeducational evaluation of their children and to present their results at hearings. 20 U.S.C. § 1415(b)(1)(A) (1978); 45 C.F.R. § 121a.503 (1977).

436. The Information subtest asks several questions allegedly representing common knowledge. On it, and on all subtests, the average scale score (derived from raw scores) is 10. A scale score of five is significantly below average.

437. Am. Psychol. Ass'n, *Revised Ethical Standards*, 8 APA MONITOR 22 (March 1977). The Code of Ethics also mandates that "psychologists remain abreast of relevant federal, state, local and agency regulations . . . concerning the conduct of their practice." *Id.*

438. 440 U.S. 301 (1979), *vacating and remanding* 560 F.2d 722 (6th Cir. 1977).

required aptitude tests. Citing a provision of the National Labor Relations Act,⁴³⁹ the union requested copies of the examinations, as well as the answers and scores of the applicants. The company supplied only sample questions, descriptive literature, and validation studies, arguing that any greater disclosure would render the test unusable, would compromise the confidentiality of the employees, and would be a breach of the psychologists' code of ethics.⁴⁴⁰ The administrative law judge who first heard the case agreed that the tests were relevant to the union's grievance but, concluding that the union officials would be unable to evaluate them, permitted access contingent on their being disclosed only through an industrial psychologist selected by the union. He also forbade the copying of the tests and their disclosure to third parties.⁴⁴¹ However, he did order that test scores be given directly to the union. On review, the NLRB affirmed the decision in principle but amended it to allow the tests and scores to be given directly to the union. The Sixth Circuit affirmed the Board's decision and dismissed contentions by the APA in an amicus brief supporting the company that disclosure of the tests would ignore the interests of the present and future examinees as well as psychologists:

[T]he principles which underlie the National Labor Relations Act are paramount in this case. Detroit Edison cannot rely upon an asserted privilege [confidentiality] which is personal to the employees who took the examination, and we are not informed as to any rule of law under which the professional code of the American Psychological Association can stand as a barrier to the right of a duly chosen and certified collective bargaining representative to receive information of use to it in carrying out its duties and responsibilities. The Board showed its consideration for the expressed concerns of the company and the Psychological Association by adopting the limitations on use of the material recommended by the administrative law judge.⁴⁴²

The Supreme Court reversed the Sixth Circuit's decision, holding that the lower court's approval of the NLRB's order requiring Detroit

439. 29 U.S.C. § 158(a)(5) (1976). This section creates a duty of employers to respond to requests by employees' bargaining agents for relevant and necessary information.

440. 440 U.S. at 307-08, 317.

441. *Id.* at 309-10.

442. 500 F.2d 722, 726-27 (6th Cir. 1977). For a critique of the decision, see Comment, *Psychological Aptitude Tests and the Duty to Supply Information*, 91 HARV. L. REV. 869 (1978).

Edison "to turn over the test battery and answer sheets directly to the Union"⁴⁴³ and "unconditionally to disclose the employee scores"⁴⁴⁴ was erroneous. The Court found that the NLRB's remedies for handling employees' grievances did not adequately protect test security:

The finding by the Board that this concern did not outweigh the Union's interest in exploring the fairness of the Company's criteria for promotion did not carry with it any suggestion that the concern itself was not legitimate and substantial. Indeed, on this record — which has established the Company's freedom under the collective contract to use aptitude tests as a criterion for promotion, the empirical validity of the tests, and the relationship between secrecy and test validity — the strength of the Company's concern has been abundantly demonstrated.⁴⁴⁵

The Court's ruling that direct disclosure of test scores to the union was inappropriate was based entirely on its desire to protect the rights of individual employees. It noted the contention that breach of the APA's code of professional ethics was an adequate ground for refusing to disclose test scores and the Board position that "the federal statutory duty to disclose relevant information cannot be defeated by the ethical standards of a private group."⁴⁴⁶ Given the importance to the employees of the confidentiality of sensitive information,⁴⁴⁷ and the minimal burden placed on the union by having to obtain consent, the Court held that Detroit Edison's resistance to disclosure of individual test results did not violate its duty to bargain in good faith.⁴⁴⁸

The application of *Detroit Edison* to school systems is limited by the fact that the decision was grounded in a specific statutory provision in a field in which the government and the courts have long been involved. Whether its rationale is applicable in the educational setting is unclear. There seems to be no barrier to parents' having access to their children's test scores under either the Buckley Amendment or Public Law 94-142. The Supreme Court's

443. 440 U.S. at 317.

444. *Id.* at 320.

445. *Id.* at 315 (footnote omitted).

446. *Id.* at 317.

447. The Court cited a number of federal and state statutes, the Buckley Amendment among them, recognizing the legislative concern for individual privacy. *Id.* at 318 n.16.

448. *Id.* at 320.

only concern in *Detroit Edison* in this regard was the lack of consent on the part of employees to the union's receipt of their test scores. Obviously, if parents request test results, there is no risk of an invasion of privacy.

The question remains whether parents may obtain the protocols. In *Lora v. Board of Education*,⁴⁴⁹ the court declared that "failure to provide parents with clinical records upon which recommendations for special . . . placement are made"⁴⁵⁰ would be a violation of due process because parents would be unable to prepare for a hearing to contest placement without them. Unfortunately, the court did not clarify what it meant by the term "clinical records." An important development in New York may have some bearing on the issue, however. Governor Carey recently signed a bill requiring testing services to make standardized admissions tests for post-secondary and professional schools available to examinees. The statute permits students to see their graded tests and their correct answers.⁴⁵¹ The law was passed despite protests from Educational

449. 456 F. Supp. 1211 (E.D.N.Y. 1978).

450. *Id.* at 1278.

451. Amendment to art. 7-A of N.Y. EDUC. LAW (July 14, 1979). See N.Y. Times, July 15, 1979, at 1, col. 1; Pecoraro, *SAT's? They're an Open Book Now in New York State*, Baltimore Sun, July 29, 1979, § K, at 2, col. 1. The bill went into effect Jan. 1, 1980. In the same month a federal district court issued a preliminary injunction, preventing the state from enforcing this law with regard to medical college entrance examinations. *Association of Am. Medical Colleges v. Carey*, No. 79-CV-730 (N.D.N.Y. Jan. 21, 1980). Cf. *Lavash v. Kountze*, 604 F.2d 103 (1st Cir. 1979) (denial of access to multiple choice employment test does not deny applicant's equal protection or due process rights).

Other so-called "Truth in Testing" legislation has been introduced in Texas, Florida, Maryland, and California. Representative Harrington introduced a bill in the 95th Congress that would have provided that any educational institution receiving federal financial assistance must provide copies of standardized tests administered to at least 10,000 individuals and the correct answers to any examinee requesting them in writing. It would have also prevented release of test scores without written permission of the examinee and required that release of test scores be accompanied by a statement indicating that they were only approximations of the true score. H.R. 6776, 95th Cong., 1st Sess. (1977). Because of problems in its language, Rep. Harrington later withdrew the bill. Two new bills have been introduced in the first session of the 96th Congress. H.R. 3564 would affect educational tests used for admission to postsecondary institutions and for employment. Each testee would be provided a written description of what the test was designed to measure, the reliability (margin of error) of the test, and the manner in which test results would be distributed. It does not explicitly provide for access to the test questions themselves. H.R. 4949 is modeled on the New York legislation. It provides that upon request test subjects may receive a copy of the test questions used in determining test scores, their answers, and the correct answers. The bill would apply to all standardized tests used for post-secondary school admissions administered to more than 5,000 test subjects each year.

Testing Service, the College Entrance Evaluation Board, and college administrators that the statute would violate test security, increase administrative costs, reduce services to handicapped persons, and disadvantage minorities. Although the statute does not affect public school tests used for special education placement, the passage of such legislation may serve as a useful precedent for access to test protocols by parents for such purposes.

CONCLUSION

Fifteen years ago two prescient scholars stated, "It requires no Cassandra to predict lawsuits by parents, and a spate of restrictive legislation, if those who administer . . . tests in schools — even for the most legitimate of scientific purposes — do not show a sensitive appreciation for both individual and group claims to a private personality."⁴⁵² Their only error was in failing to divine all the causes for the litigation and legislation that they so accurately predicted. While concern for privacy did stimulate such suits as *Merriken* and such statutes as the Buckley Amendment, it was the perversion of psychological tests as instruments of segregation and discrimination that led to the case and statutory law that this Article has surveyed. The ineluctable conclusion is that psychological testing — now banned, condemned, and regulated — would have survived with its reputation relatively unscathed had it not been used both intentionally and inadvertently to perpetuate racial separation, stereotyping, and stigmatization.

An often repeated rejoinder to criticisms of psychological testing has been that it is not the instruments but the users who have brought about the current state of affairs. While such a defense serves to deflect attention from genuine flaws in the tests themselves, there is truth in it. It is the use of tests that courts enjoin, but it is the testers and their employers who are sued. It was the failure of psychologists to question their role, to scrutinize the psychometric soundness of their instruments, and to test the validity of their interpretations that resulted in misclassification and miseducation and the injuries that flowed therefrom to significant numbers of children of all racial and cultural backgrounds. The consequence has been the imposition of a number of well-meaning but unrealistic and often impossible restrictions on the use of psychological tests.

There are at least three benefits, however, from the increased involvement of courts and legislatures in psychologists' testing

452. Reubhausen & Brim, *Privacy and Behavioral Research*, 65 COLUM. L. REV. 1184, 1194 (1965).

practices. First, it has made the profession, as well as society generally, more sensitive to racial and cultural differences and to how apparently innocent and benign practices may perpetuate discrimination. Second, it has alerted psychologists and other mental health professionals to the fact that they will be held responsible for their conduct. Because of the accountability mechanisms now inherent in the procedural protections afforded handicapped children and their parents, psychologists who work in school settings find that they cannot view themselves only as passive recipients of orders from their supervisors. To protect the rights of their clients, to safeguard their own integrity, and, in the long run, to serve the asserted ends of their employers to educate students effectively, they must examine their practices, their interpretations, and their ultimate recommendations. Finally, the attack on psychological testing has accelerated the search for alternative means of assessment so that what is said about children is a more valid, truer depiction of how they perceive themselves and how they function in all spheres of life. In this light, the intense and searching examination that psychological testing has received from the legal system should be viewed as both salutary and welcome.